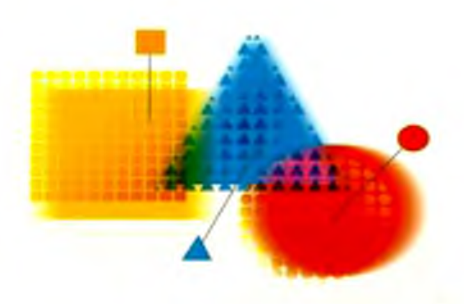
Alicia R. W. de Camilloni, Susana Celman, Edith Litwin y María del Carmen Palou de Maté

La evaluación de los aprendizajes en el debate didáctico contemporáneo

Paidós Educador



Alicia R. W. de Camilloni, Susana Celman,
Edith Litwin y M. del Carmen Palou de Maté

La evaluación de los aprendizajes en el debate didáctico contemporáneo



PAIDÓS

Buenos Aires
Barcelona
México

Cubierta: Gustavo Macri

371.26 LAE

La evaluación de los aprendizajes en el debate didáctico contemporáneo / Alicia R. W. de Camilloni... [et al.].- 1ª ed. 6ª reimp.- Buenos Aires : Paidós, 2008.
176 p. ; 22x16 cm.- (Educador)

ISBN 978-950-12-2129-9

I. De Camilloni, Alicia R. W. 1. Aprendizaje-Evaluación

1ª edición, 1998
6ª reimpresión, 2008

Queda hecho el depósito que previene la Ley 11.723
Impreso en Argentina - Printed in Argentina

Impreso en Primera Clase,
California 1231, Ciudad de Buenos Aires, en julio de 2008
Tirada: 1000 ejemplares

ISBN 978-950-12-2129-9

SUMARIO

PRESENTACIÓN

En el debate didáctico contemporáneo, la evaluación constituye un tema que inevitablemente se plantea como de difícil resolución. Este libro reúne una serie de trabajos que intentan dar respuestas a la problemática evaluativa, sin dejar de ahondar en los orígenes de la dificultad de su resolución. Estas reflexiones inscritas en la didáctica contemplan el análisis de la evaluación incorporado a los procesos de enseñar y aprender. Evaluar al margen de esos procesos y de la institución en la que se enmarcan nos remitiría a una visión instrumentalista que probablemente no daría cuenta del aprendizaje construido. La evaluación como campo y herramienta de conocimiento, la evaluación de las prácticas docentes, los procesos de autoevaluación, la evaluación en las instituciones, los criterios para construir evaluaciones y los sistemas de calificación son los temas que se desarrollan para ofrecer respuestas a este problema didáctico. Los capítulos fueron escritos por profesoras investigadoras de la Universidad de Buenos Aires, Entre Ríos y Comahue y forman parte del diálogo que los didactas de estas universidades públicas sostienen asiduamente con el objeto de pensar cooperativamente los problemas de la educación, de la enseñanza y del aprender.

E. L.

LA EVALUACIÓN: CAMPO DE CONTROVERSIAS Y PARADOJAS O UN NUEVO LUGAR PARA LA BUENA ENSEÑANZA

Edith Litwin

Para Ángeles y Marisa

El campo de la evaluación da cuenta de posiciones controvertidas y polémicas no sólo desde una perspectiva política, sino también desde la pedagógica y didáctica. Este artículo intentará abordar algunas de las controversias en este último terreno, el de la didáctica, presentándolas a partir de un análisis crítico y de las nuevas derivaciones de los enfoques cognitivos. Inscribiremos el análisis de estas discusiones en el estudio de la evaluación en relación con las prácticas de los docentes y las implicancias en el aprendizaje de los estudiantes. Nos interesa desarrollar algunas reflexiones en torno a la evaluación desde nuestra concepción de una nueva agenda de la didáctica, lo cual implica inscribirla en las derivaciones de la psicología cognitiva, fundamentalmente, en la clase reflexiva, en la comunicación didáctica y desde la problemática moral que entraña. Por último, incluimos el análisis de un caso que nos permite reconocer las categorías descritas y señalar los problemas con los que hoy nos enfrentamos.

LA EVALUACIÓN COMO CAMPO DE CONTROVERSIAS: UNA CUESTIÓN DIDÁCTICA –TEMA CENTRAL, PERIFÉRICO O DE BORDE DE LA AGENDA–

Reconocemos a la evaluación integrando el campo de la didáctica, definida como teoría acerca de las prácticas de la enseñanza. En dicho campo, algunas dimensiones tales como el contenido y el método han sido centrales en tanto las ideas y la prácticas de la

enseñanza se han configurado en torno a ellas. Otras cuestiones como el currículo, las estrategias o la evaluación, si bien formaron parte de la agenda clásica de la didáctica y se consolidaron como categorías del debate didáctico, no tuvieron ese rango de centralidad. Sin embargo, el problema de la evaluación, ha ido adoptando progresivamente una mayor importancia. Pero dicha importancia fue adquirida como resultado de una "patología": muchas prácticas se fueron estructurando en función de la evaluación, transformándose ésta en el estimulo más importante para el aprendizaje. De esta manera, el docente comenzó a enseñar aquello que iba a evaluar y los estudiantes aprendían porque el tema o problema formaba una parte sustantiva de las evaluaciones.

Díaz Barriga analiza estas cuestiones refiriéndose a la evaluación como el lugar de las inversiones (Díaz Barriga, 1993). Es el examen, según este autor, el instrumento que permite invertir los problemas sociales en pedagógicos y el que favorece que el debate educativo deje de ser un problema conceptual para convertirse en un problema técnico. Son las inversiones las que traicionan la función que se le pretende asignar al instrumento.

Nosotros sostenemos, además, que en las prácticas de enseñanza, la actitud evaluadora invierte el interés de conocer por el interés por aprobar en tanto se estudia para aprobar y no para aprender. Es el mismo profesor que, cuando enseña un tema central o importante de su campo, destaca su importancia diciendo que será evaluado y lentamente va estructurando toda la situación de enseñanza por la próxima situación de evaluación. Sin embargo, estos debates acerca de la centralidad como patología podrían modificarse si los docentes recuperan el lugar de la evaluación como el lugar que genera información respecto de la calidad de su propuesta de enseñanza. Desde esta perspectiva, la evaluación sería tema periférico para informar respecto de los aprendizajes de los estudiantes, pero central para que el docente pueda recapacitar respecto de su propuesta de enseñanza. Lo patológico sería considerarlo como el lugar de información indiscutible respecto de los aprendizajes de los estudiantes y no así como un lugar privilegiado para generar consideraciones de valor respecto de la propuesta metódica y los procesos del enseñar de los docentes. Es también patológico que la evaluación estructure las actividades docentes.

Las paradojas de la evaluación de las prácticas de enseñanza o de la supervisión docente

Apreciar, estimar, atribuir valor o juzgar han sido los conceptos que más se asociaron a la evaluación. Desde una perspectiva didáctica, el concepto implica juzgar la enseñanza y juzgar el aprendizaje; atribuirles un valor a los actos y las prácticas de los docentes y atribuirles un valor a los actos que dan cuenta de los procesos de aprendizaje de los estudiantes.

Desde una perspectiva didáctica, significa también el estudio de las relaciones y de las implicancias del enseñar y aprender. En esta situación bipolar y de mutuas relaciones, tales acciones en el contexto de la didáctica tuvieron significaciones diferentes. La evaluación de las prácticas fue planteada como un tema de supervisión docente y tuvo la característica del control de la actividad. No se la planteó como metodología para el mejoramiento de las prácticas, sino que adquirió un lugar de reconocimiento de su existencia. Instancias muy formales o totalmente formalizadas a través de un instrumento permitieron acreditar o reconocer la existencia de dichas prácticas para la continuidad del ejercicio docente.

Por otra parte, diversas metodologías para el análisis de las prácticas, estudios sobre el pensamiento del docente y su relación con ellas, dan hoy cuenta de que la evaluación se constituye en un tema de interés. Sin embargo, se lo aborda como nuevo objeto problemático desde la metodología de investigación y fundamentalmente como reconocimiento de un tema que puede ser objeto de estudio en tanto da cuenta de categorías que estaban soslayadas en el análisis de las prácticas docentes. Con esto queremos significar que la evaluación de las prácticas como problema didáctico es casi inexistente. Sin embargo, en los procesos de formación docente, las prácticas experimentales o residencias como procesos formativos plantearon un lugar para su evaluación, en general, asociados a los procesos preactivos, esto es, al diseño de las prácticas. Pero al hacerlo sus implicancias fueron las mismas de la evaluación de los aprendizajes de los alumnos –en estos casos, los alumnos residentes o practicantes– y nunca asociadas al lugar profesional de la práctica.

Algunas contradicciones en la evaluación de los aprendizajes

En relación con la evaluación de los aprendizajes como campo y problema, ésta siempre estuvo relacionada con procesos de medición de los mismos, la acreditación o la certificación, y rara vez con el proceso de toma de conciencia de los aprendizajes adquiridos o con las dificultades de la adquisición, de la comprensión o la transferencia de algunos temas o problemas.

La suposición de la que parten estas evaluaciones consiste en señalar que es posible medir los aprendizajes en el mismo momento que ocurren o dentro de un curso escolar y que, para ello, es útil recurrir a la creación de situaciones más formalizadas en las que la medición puede ser realizada. Esta situación es controvertida para nosotros en tanto reconocemos que los aprendizajes significativos necesitan tiempos de consolidación en los que los temas o problemas enseñados pueden ser relacionados con otros y que, probablemente, el verdadero aprendizaje tenga lugar cuando el alumno o la alumna se encuentren fuera del sistema o de la clase en donde se planteó la situación de enseñanza.

En más de una oportunidad la evaluación, además, plantea una situación de sorpresa con el objeto de que el alumno no se prepare para ella y así poder reconocer fehacientemente los aprendizajes. Son característicos de estas situaciones el estrés o el shock admitidos como la situación habitual de los exámenes. Estas prácticas no tienen por objeto reconocer lo aprendido, sino analizar lo consolidado o lo que resiste pese a la presión o en una situación de tensión. Evidentemente, esto es ajeno a un ambiente natural en donde se reconozca un aprendizaje o se favorezca su concreción.

También, en la búsqueda de reconocer verdaderos procesos de transferencia, muchas veces las evaluaciones implican exigencias de procesos reflexivos novedosos que nunca formaron parte de los procesos de enseñanza. La evaluación no mejora lo aprendido, sino que permite, en el mejor de los casos, su reflejo.

Otra de las cuestiones controvertidas de las prácticas evaluativas es el papel que cumplen las ideas que los docentes tienen acerca de los aprendizajes de los alumnos. Un estudio clásico de la literatura pedagógica como el de Rosenthal y Jacobson describió el papel de las creencias de los docentes en los procesos del aprender de los

estudiantes (Rosenthal y Jacobson, 1980). En el trabajo de investigación que realizaron mostraron cómo las expectativas que los docentes tienen acerca de los posibles rendimientos de sus alumnos pueden convertirse en una profecía que se cumple inexorablemente. Buenos y malos alumnos desde las representaciones de los docentes generan buenos y malos rendimientos. En nuestra experiencia, la instalación de un alumno en bueno o malo por parte de docentes, autoridades escolares, padres y desde el mismo alumno genera actitudes hacia el estudio, cumplimiento y esfuerzo en el mismo sentido de la instalación. En un estudio de las calificaciones que los docentes adjudican a sus alumnos, que hemos realizado durante más de cuatro años y con grupos diferentes, observamos una clara tendencia en la mayoría de los docentes por mantener los mismos porcentajes de aprobación o desaprobación de los estudiantes, independientemente de las diferencias de las cohortes. Por otra parte, el halo que rodea a algunas de las conductas de los estudiantes, como el orden, la buena letra, la participación en clase, ha llevado a que se consignen estas actitudes o actividades como expresión de los buenos aprendizajes y tiene influencia en las evaluaciones que realizan los docentes, incrementando o disminuyendo las calificaciones según los casos. No queremos decir con esto que estas conductas no puedan estar asociadas a los aprendizajes, sino que no necesariamente están asociadas a la enseñanza.

La evaluación: un análisis desde algunas derivaciones de la psicología cognitiva

En un trabajo anterior (Litwin, 1997: 84-95), al reconocer nuevas perspectivas de análisis en la agenda de la didáctica, hemos estudiado principios que fomentan la comprensión, abordajes para la enseñanza que emergen de la investigación cognitiva, disposiciones del buen pensar y características de una enseñanza que fomenta la reflexión y el pensamiento crítico. Analizar la evaluación desde esta perspectiva significa reconocer las posibles maneras de comprender de los estudiantes, tanto por parte de los docentes como de los propios alumnos. Ese proceso de reconocimiento conformaría

también un proceso de conocimiento que se integraría a la situación de enseñanza.

Nuestra primera reflexión desde este marco, entonces, consiste en considerar que la evaluación es parte del proceso didáctico e implica para los estudiantes una toma de conciencia de los aprendizajes adquiridos y, para los docentes, una interpretación de las implicancias de la enseñanza en esos aprendizajes. En este sentido, la evaluación no es una última etapa ni es un proceso permanente. El lugar propicio tiene que ver con el lugar de la producción de conocimiento y la posibilidad, por parte de los docentes, de generar inferencias válidas respecto de este proceso. Carece de sentido la adquisición de una actitud evaluativa constante porque no permitiría desarrollar situaciones naturales de conocimiento o intereses no suscritos en una temática directamente involucrada en el aprendizaje de una materia o tema. Se desvirtuaría de esta manera el sentido del conocimiento al transformar las prácticas en una constante de evaluación. Por otra parte, las propuestas de construcción del conocimiento en las situaciones que lo posibilitan, dan cuenta de las situaciones de transferencia, permiten pensar bien los problemas del campo, tratan de resolver problemas reales y se constituyen, en la medida de lo posible, en procesos de resolución de problemas genuinos y consistentes con el campo de conocimientos de que se trata.

Desde esta perspectiva, tanto la identificación como la búsqueda de problemas o la resolución de los interrogantes genuinos constituyen los mejores desafíos a la hora de comprender un campo de conocimiento. Por otra parte, si los procesos de análisis, síntesis, aplicación de principios –comparativos o de generalización– están implicados en los procesos de producción, debieran ser identificados porque dicho reconocimiento favorece la resolución de la producción. En algunas oportunidades solicitamos la utilización de una información adquirida en otros contextos sin reconocer que en la base de la actividad estamos involucrando procesos que implican además de la descontextualización, la abstracción y la aplicación, la memoria y la recuperación de información.

A la hora de reflexionar sobre la evaluación, sostenemos los mismos interrogantes que a la hora de pensar las actividades y su valor en la construcción del conocimiento. Una actividad "pedago-

gizada", esto es, una actividad inventada para la enseñanza, carente de valor y sentido real, no implica ningún atractivo o desafío para los estudiantes, no los compromete ni social ni cognitivamente y cobra un valor totalmente extrínseco: actividad para ser evaluada. Cuando la actividad implica un desafío genuino es más sencillo adoptar una propuesta evaluativa porque su mismo carácter de indefinición y de compromiso permite reconocer verdaderos retos cognitivos.

En nuestra experiencia, los alumnos acumulan a lo largo del sistema educativo variadas propuestas de reproducción de los conocimientos, en donde el almacenamiento de la información juega un lugar privilegiado. Evaluar el almacenamiento de información, esto es, la memoria reconocida a largo o a corto plazo en situaciones en donde el alumno fundamentalmente recuerda hechos y datos, ha sido una práctica constante en los diferentes niveles del sistema educativo. Desde una perspectiva cognitiva, planteamos actividades que cambien el lugar de la evaluación como reproducción de conocimientos por el de la evaluación como producción, pero a lo largo de diferentes momentos del proceso educativo y no como etapa final. Sostener que existen diferentes momentos implica diferenciarlos y reconocer que algunos son mejores que otros, pero no todos son momentos propicios para la evaluación. El buen sentido, el sentido pedagógico, nos permite superar la fiebre evaluadora. Nuestra propuesta consiste en reconocer los momentos clave en donde una buena información acerca de las características del aprender nos ayuda a mejorar el mismo aprendizaje, focalizando los problemas, las dificultades o los hallazgos.

Quizás el principal problema reside en la construcción de los criterios con que se evalúan las actividades. Éstos facilitan los juicios y permiten el mejoramiento de las prácticas o producciones que nos hayamos propuesto. Para cada actividad es posible que, como docentes, nos planteemos cuáles son los criterios que nos permiten reconocer su concreción. Los criterios son recursos muy potentes para evaluar las producciones de los estudiantes, pero son elaborados a partir de las experiencias y, por tanto, no son infalibles ni debieran cristalizarse. Son instrumentos que nos ayudan a reconocer el valor de las actividades.

Matthew Lipman (1997) desarrolla en este sentido una serie de propuestas y debates que generan derivaciones para las prácticas

evaluativas. Por ejemplo, sostiene como criterios la representatividad, la significación y la diferenciación cognitiva. Desde la representatividad, en un trabajo debiera reconocerse la implicación del tema tratado en el conjunto de los temas o cuestiones del campo; la significación refiere a advertir el valor o importancia del tema; y la diferenciación cognitiva considera el proceso reflexivo implicado: memoria, traducción, análisis, síntesis, resolución de problemas, etcétera. Los criterios que se debieran explicitar a los estudiantes permiten al docente emitir un juicio acerca del valor de la tarea desempeñada. A su vez generan una corrección globalizada, integral, más acorde con la significación de una tarea compleja. Entendemos el valor de explicitar los criterios a los estudiantes y, en especial, aquellos que consideramos implícitos y sobre los que se generan múltiples malentendidos. Por ejemplo, la presentación, la pulcritud o la ortografía suelen estar implícitos para el docente pero no así para nuestros estudiantes. Por otra parte, reconocer estos criterios y su valor en el momento de la evaluación nos permite aceptar el lugar que ocupan y no utilizarlos solamente para refirmar nuestras ideas frente a las valoraciones de los aprendizajes de algunos estudiantes.

Una importante derivación cognitiva a la hora de pensar en las evaluaciones conlleva una situación paradójica: se reconoce pormenorizadamente el proceso reflexivo implicado, pero se evalúa global e integralmente la tarea realizada. A la hora de seleccionar los criterios para evaluar las actividades, producciones o posibilidades de interrogarse de los estudiantes, también importa la construcción de megacriterios (Lipman, 1997), que no son otra cosa que criterios, pero que, de alguna manera, son relevantes como tales; su trayectoria en el pasado da cuenta de que son confiables para emitir juicios y tienen fuerza o potencia comparados con otros. La consulta bibliográfica, un aporte original o la pulcritud en el análisis pueden constituir, según los casos, megacriterios en la evaluación. La índole de la tarea en relación con el campo de conocimiento es la que favorece la concreción de los criterios a partir de los que reconocemos su valor.

La evaluación y la reflexión en la clase

En un ambiente en el que se privilegia el pensar, rara vez nos preocupamos por la evaluación. Sin embargo, evaluar esta actividad privilegiada de la naturaleza humana nos propondría reconocer el interés de la escuela por suscitar este tipo de actividades. En el marco en el que las actividades reflexivas se producen reconocemos que el mundo es ambiguo e inequívoco, las disciplinas no representan el total del conocimiento y a menudo se yuxtaponen, el docente es falible y la mejor expresión del conocimiento es el razonamiento del estudiante acerca de un tema o cuestión.

Jerome Bruner (1997), en una de sus más recientes publicaciones, nos plantea nueve postulados que pueden guiar una aproximación a la problemática educativa desde una perspectiva psicocultural, esto es, tanto desde el lugar en el que los estudiantes construyen realidades y significados en sus subjetividades hasta un lugar macro en el que se inscriben los sistemas de valores, derechos, intercambios y poderes de una cultura. Estos postulados, tal como los plantea Bruner, nos permiten reconocer aproximaciones diferentes de la problemática educativa, en tanto la educación es una búsqueda compleja por adaptar una cultura a las necesidades de sus miembros y, al mismo tiempo, por adaptar a sus miembros y sus modos de conocer a las necesidades de la cultura. Hemos construido a partir de estos postulados una serie de implicancias para la evaluación que pueden orientar o señalar las dificultades que se generan cuando pretendemos evaluar.

En primer lugar, Bruner cita el postulado perspectivista, por el que reconoce que la construcción de cualquier hecho o proposición es relativa al punto de vista desde el cual el término es construido. Comprender algo de determinada manera no excluye la comprensión desde otro camino. Esto no impide reconocer la lógica de la construcción, el sentido común o las diferencias posibles entre alternativas por la coherencia con otros marcos de pensamiento, evidencias, etcétera. La evaluación, desde este postulado, alentará la comprensión de los caminos alternativos para la construcción de conocimientos y la erradicación de "verdadero o falso" desde una perspectiva particular.

El postulado de los límites reconoce que, en tanto humanos, nos

hemos especializado en ciertos modos de conocer, pensar, sentir o percibir y, por otra parte, que las experiencias previas afectan lo que hoy estamos pensando. Si atendemos a este principio, deberíamos reconocer que es fundamental generar propuestas que vayan más allá de las predisposiciones naturales y, por lo tanto, valorar la trasmisión de todas las herramientas culturales que lo posibilitan. Los sistemas simbólicos constituyen herramientas privilegiadas para lograrlo, en tanto mejoran la capacidad de construcción de significados. Reconocer su utilización y analizar cómo construimos el conocimiento pareciera una de las conclusiones más significativas de esta propuesta. Elliot Eisner (Eisner, 1997), en una expresiva metáfora, sostiene que si la única herramienta de que disponemos es un martillo, veremos al mundo y sus objetos como si fueran clavos.

El postulado del constructivismo explicita implicaciones del postulado anterior, en tanto reconoce la realidad como una construcción social. La educación es una ayuda para que los niños aprendan a usar las herramientas que permiten la creación de sentido y significado. Adaptarse al mundo, entenderlo y aprender a cambiarlo conforman la propuesta de la educación. Reconocer, mediante la evaluación, la comprensión de la realidad y la comprensión crítica, se constituye en otro acto de conocimiento y, por tanto, de construcción.

El postulado interactivo recupera la relación con el otro en el proceso de construcción del conocimiento. Evidentemente este principio da cuenta de las dificultades de generar un proceso evaluativo. El compromiso y la participación "del otro", docente o alumno, y sus implicaciones impiden separaciones taxativas a la hora de evaluar. En más de una oportunidad, la participación del docente con una palabra o contextualizando el problema o tema, permite una resolución. "Soplar" suele ser considerado como una situación de fraude cuando en realidad es una situación de ayuda. El problema reside en que "soplar" en una situación de evaluación, en la que prima la consideración del almacenamiento de información y recuperación, invalida la misma actividad. En cambio, si la actividad es la resolución de un problema, la ayuda que proporciona ofrecer una palabra debe ser contemplada como parte de la resolución de la misma actividad.

El postulado de externalización nos remite al reconocimiento de

que aprender implica realizar ensayos, resolver problemas, plantear preguntas. Las producciones implican auténticos desafíos y permiten reconocer las dificultades, limitaciones y potencialidades del proceso del aprender. Bruner, al analizar este principio, señala que la externalidad propone un récord para nuestros esfuerzos mentales en donde las producciones grupales pueden ayudar a fomentar la existencia de comunidades en las que las obras colectivas permiten redes solidarias y, en tanto productos, favorecen los procesos reflexivos y la reflexión acerca de cómo se llevaron a cabo.

El postulado instrumentalista reconoce que en las escuelas siempre se ha seleccionado un uso particular de la mente. La selección tiene que ver con las necesidades de la sociedad respecto de los individuos que la conforman. Se distribuyen habilidades en forma diferenciada que dan cuenta de las implicancias políticas del accionar educativo y de las consecuencias sociales y económicas del acto de educar. Nos preguntamos, en este marco, si la evaluación ocupará el lugar de legitimación de las políticas, permitirá dar cuenta de las contradicciones que los talentos y las individualidades generan en términos políticos o podrá transparentar y romper el sentido instrumentalista del accionar educativo.

El postulado institucional requiere entender la escuela como institución con sus lógicas y roles diferenciados, entre otras cuestiones. Aprender el oficio de alumno que implica el conocimiento de las rutinas y ritualidades, dejarse evaluar por otros y asentir en esa evaluación como una de las prácticas más comunes. Es difícil recuperar el buen sentido, el sentido pedagógico ante tantas prácticas que institucionalizaron la evaluación desde el lugar de la administración del sistema, del control, del reaseguro de su supervivencia. Desde este principio, la evaluación debiera contradecir las propuestas habituales para plantear en el marco de la misma institución un nuevo proyecto escolar que recupere el sentido de la enseñanza y no la enseñanza desde su medición con el objeto de acreditar o certificar.

El postulado de identidad y autoestima implicado, según Bruner, en todos los demás, cobra especial significación en el campo de la evaluación, en tanto los éxitos y los fracasos son las nutrientes principales en el desarrollo de la personalidad. En la escuela se aplican criterios que, en muchas oportunidades, arbitrariamente

condicionan los éxitos y los fracasos. Por tanto, el lugar de la autoestima es comúnmente vapuleado en las prácticas evaluativas. En nuestras historias educativas muchos evocamos situaciones de injusticia que generaron recuerdos imborrables. Quizás esto explique, sin justificarlo, el fraude que tan corrientemente se encuentra asociado a la evaluación.

El último postulado planteado por Bruner es el narrativo. En nuestra cultura tecnologizada reconocemos, sin subvalorar el pensamiento lógico-científico, la importancia de la narrativa tanto para la cohesión de una cultura como para estructurar la vida de un individuo. Bruner señala la existencia de dos modos de pensamiento: el lógico-científico y el narrativo. La escuela siempre dio preponderancia al primero. Quizás encontremos en ello las razones de tantas dificultades de comprensión. En nuestra investigación didáctica hemos reconocido, en el discurso del docente, variaciones en el arte de narrar que posibilitaron los procesos de comprensión. Estas variaciones, asociadas al tratamiento de los contenidos, provocan en algunos casos mayor participación y permiten reconocer, a la hora de la evaluación, las implicancias reflexivas. Las clases en las que su narrativa se organiza alrededor de un objeto, experimento, caso u obra paradigmática, favorecen la participación del alumno. Se trata, en general, de situaciones que se expresaron desde una narrativa lineal, atractiva y fácilmente comprensible y que luego fueron analizadas con todas sus implicancias en sus respectivas dimensiones de análisis. Los hechos, datos, hipótesis son reconocidos cuando se los reinterpreta a la luz de una narración poderosa que se desarrolló en el primer momento de la clase. El interés de los estudiantes se mantiene vivo y se expresa en sus preguntas, ejemplos y argumentaciones.

El conjunto de estos postulados referidos a la enseñanza, en el que desarrollamos algunas reflexiones para la evaluación, constituye una manera de pensar los problemas de estas prácticas. Instala a la evaluación como un conocimiento que se construye en una institución en la que las prácticas cobran sentido político y en donde los espacios de reflexión tienen un sentido privilegiado.

Desde esta perspectiva, las evaluaciones que vuelven a los alumnos con nuestros aportes y sugerencias para permitir que continúen en el proceso de producción dan cuenta de las posibilidades de

aprender, tanto de los docentes, en lo que respecta a que encuentren los lugares de mejoramiento o crecimiento del trabajo, como las de los alumnos, al favorecer el desarrollo de sus producciones. Se llevarán a cabo cuantas veces sea posible y finalizarán en el momento en que así lo acuerden, compartidamente, docente y alumno. Las distintas evaluaciones, además, tratarán de reconocer los diferentes modos de representación en que se pueden expresar los procesos de producción de conocimiento. Ensayos, elaboración de proyectos, análisis de fuentes, resolución de casos, representaciones, críticas a una producción, pueden dar lugar a elaboraciones de distinto tipo atendiendo a las diversas maneras de expresión de los estudiantes.

En síntesis, en el acto de evaluar nos preocupan el reconocimiento de los límites, autolímites e imposiciones en el conocer, las prácticas que se configuran en desmedro de la autoestima y la organización de la misma como actividad solitaria o cooperativa. Nos interesa reconocer el valor de elaborar y explicitar los criterios que utilizaremos en las prácticas evaluativas, reconocer nuestros límites como docentes (en tanto configuramos en nuestras clases una auténtica labor de construcción de conocimiento) y proponer formas y propuestas que contemplen la diversidad de las expresiones del saber.

La evaluación y la comunicación didáctica

Al analizar a la evaluación desde esta perspectiva reconocemos en primer lugar dos aspectos ficcionales de la comunicación educativa. El primero consiste en que el o la docente entabla el diálogo con los alumnos para indagar lo que éstos desconocen e intentan ocultar. El segundo se refiere al carácter ficcional de la pregunta que hace el docente, en tanto sólo la efectúa cuando sabe la respuesta. Evidentemente, en una auténtica conversación no se pregunta lo que se sabe ni se indaga para descubrir el desconocimiento. Animarse a preguntar, por parte de los alumnos, también implica atreverse a descubrir lo que se ignora.

No obstante, en la conversación en clase podemos reconocer múltiples prácticas evaluativas que favorecen los procesos del pensar, siempre que el clima del aula sea favorable a estos procesos y no

se encuentre sumergido en una práctica de medición. Con esto queremos señalar que la ficción en la comunicación es una de las tantas expresiones de los ritos en las clases, mientras que las auténticas preguntas se configuran como el espacio privilegiado para la construcción del conocimiento. Cuando el maestro se constituye en un interlocutor más sabio y con más experiencia, permite que los estudiantes planteen sus hipótesis más arriesgadas, sus intuiciones o sus interrogantes sin temer que, con estas intervenciones, el docente descubra las ignorancias. Bruner plantea (Linaza, 1996), que le gustaría poder describir la institución educativa como una especie de área de seguridad que le permite a uno ser más atrevido, tener hipótesis o ponerlas en práctica. Curiosamente, señala, no siempre ha resultado así. Ésta busca evaluar a las personas, determinar quién logra notable, sobresaliente o matrícula, y quién se queda sólo con un aprobado. Esto supone, afirma Bruner, clasificar a la sociedad, lo que provoca una actitud cautelosa por parte de las personas que asisten a ella, en lugar de estimular la "zona de desarrollo próximo", tal como la denominaba Vigotsky.

Las preguntas de los alumnos le permiten al docente reconocer las maneras en que aquéllos se interrogan respecto de un campo, la naturaleza de los errores o las falsas concepciones. En una clase, favorecer los espacios para que los alumnos se interroguen e interroguen, aun cuando esto corte la exposición del docente, implica promover el pensar. Responder a las preguntas de los estudiantes, por otra parte, implica incorporar la pregunta en el discurso y ofrecer una respuesta que dé continuidad también al discurso.

Cuando el docente pregunta, más de una vez se encuentra con que el alumno responde rápidamente, sin detenerse a pensar. En esos casos, es conveniente solicitarle al alumno que se tome tiempo para pensar. Quizá nos explicamos esta rápida respuesta de los estudiantes por las prácticas habituales de ensayo y error que generan un residuo cognitivo, esto es, una habilidad cognitiva nueva dada por el estímulo de nuestra cultura a esta manera de pensar. Paradójicamente, nos encontramos con que, en las situaciones de exámenes orales, los docentes realizan preguntas que sólo se pueden contestar con una respuesta rápida y, por tanto, conducente a esta situación de ensayo y error. Los tiempos de la maduración de

una idea o simplemente de reflexión más de una vez son tiempos diferentes a las prácticas de acreditación o certificación.

También los docentes suelen pensar que no evalúan pormenorizadamente, sino de manera holística o integral, y que para aprobar una materia o campo no se requiere un nivel especializado, sino un mínimo de conocimientos que, en realidad, refiere a una cultura general. Sin embargo, y contradictoriamente, estos conocimientos generales son los únicos que desdeñan enseñar.

La evaluación desde una perspectiva moral

El prejuicio, el evaluar para corroborar la calificación que anticipamos, el descalificar, esto es, el disminuir la calificación a partir del reconocimiento de errores, sean de índole temática, de ortografía o de presentación, no parecen constituir buenas prácticas evaluativas. Los análisis que hasta aquí hemos propuesto nos han permitido reconocer tanto las contradicciones o las paradojas de las prácticas evaluativas como algunas dimensiones de análisis que favorecen la construcción de buenas propuestas para la evaluación. Entendemos que el análisis de un caso nos permitiría nuevas reflexiones en el marco de esta dimensión moral que deseamos considerar.

La evaluación de los aprendizajes en un curso de la universidad o cómo favorecer la construcción del conocimiento acerca de los aprendizajes

Esta experiencia la llevamos a cabo en la última evaluación del curso de una materia de la carrera de Ciencias de la Educación, en la Facultad de Filosofía y Letras de la Universidad de Buenos Aires, con el objetivo de posibilitar el reconocimiento de los alumnos de sus propios aprendizajes. Se realizó en jornadas diferentes. La primera implicó aproximadamente tres horas de trabajo. La segunda se concretó a partir del trabajo individual de tres docentes y de una reunión del mismo grupo de aproximadamente tres horas de duración. Las etapas siguientes se realizaron durante cuatro horas de trabajo con la participación de los docentes y los alumnos.

Primera etapa

Se les solicitó a los estudiantes que en forma individual confeccionaran un parcial. Los criterios que se plantearon para la confección fue que éste no refiriera a preguntas memorísticas, sino que su realización significara, para el alumno que lo iba a responder, un trabajo complejo que intentara reflejar los problemas del campo y que permitiera un trabajo de elaboración. Junto con los alumnos, un docente de la cátedra también participó de esta tarea de elaboración de la evaluación.

Una vez confeccionados los parciales, se pidió a los alumnos que se intercambiaran las propuestas y cada alumno, al igual que el docente que había formulado el parcial, respondió a uno distinto del que había confeccionado.

Segunda etapa

Los docentes de la cátedra corrigieron los exámenes, realizaron análisis de los materiales y adjudicaron una nota a cada examen. El grupo de docentes intercambió opiniones respecto de algunas resoluciones y acordó criterios para la corrección.

Las etapas siguientes tuvieron por objetivo generar en los alumnos una situación que les permitiera reconocer y valorar sus aprendizajes para obtener la acreditación.

Tercera etapa

Se leyó en voz alta y para todo el grupo el parcial que formuló el docente. Se comparó la formulación del experto con la de los alumnos. Se analizó cada pregunta o sugerencia que formuló. Se reconoció lo que subyace al material en tanto enfoque de producción y el tipo de resolución que plantea. Se analizó lo que implica de conocimiento experto la misma formulación y cómo favorece una buena resolución. A continuación se leyó el parcial que el docente contestó y se reconocieron los conocimientos profesionales o de otros campos que tuvo que utilizar para poder dar su respuesta. De esta manera se analizaron los condicionantes de las preguntas, las

respuestas que no se podían considerar porque eran respuestas posibles desde el conocimiento experto y las dificultades que plantearon los exámenes de los alumnos para su resolución. Este análisis le permitió a cada alumno iniciar su proceso de reflexión respecto de los interrogantes o la propuesta que plantearon y la respuesta que dieron. Pero no simplemente comparando su trabajo con el del experto, sino reconociendo las dificultades que tuvieron que sortear al preguntarse y al responder.

Cuarta etapa

Cada alumno analizó el parcial que formuló. Reconoció en la formulación cuáles ideas le hicieron plantear las preguntas y cómo permitieron sus ideas las resoluciones. Asimismo, diferenció las resoluciones del compañero de las que él pensó como posibles y analizó las correcciones, sugerencias y calificación que los integrantes de la cátedra escribieron frente a la respuesta. Los estudiantes utilizaron, por tanto, la respuesta al parcial que formularon y las correcciones de los expertos para autoadjudicarse una calificación.

Quinta etapa

Los estudiantes analizaron el examen que contestaron y las correcciones, sugerencias, comentarios y calificación que la cátedra atribuyó a su parcial. Esta nota no fue promediada con la anterior, sino que se constituyó en un nuevo elemento significativo para que los estudiantes pudieran ratificar o modificar la primera calificación que se adjudicaron.

Análisis del caso

En primer lugar nos cabe analizar las razones por las que incluimos un docente que realice las mismas actividades de los alumnos en todas las situaciones de examen. Entendemos que como docentes siempre tenemos que resolver el examen que formulamos. Sólo así reconocemos las dificultades que demanda la resolución y no planteamos lo que no puede ser resuelto. Por otra parte, el experto puede reconocer en la situación de evaluación si lo que permite la resolu-

ción son conocimientos propios del campo o tuvo que apelar a saberes de otros campos no desarrollados en la materia.

Si los alumnos reciben el examen en una situación de novedad, un docente de la cátedra debe también desconocer el examen y resolverlo en la misma circunstancia que el alumno. Aun cuando entendemos que en el docente no se ponen en juego las mismas situaciones de tensión en tanto no se lo califica y no acredita a través de este trabajo, mostrar su producción en una situación similar a la de los estudiantes transparenta su lugar de experto. Esta circunstancia se plantea a partir del reconocimiento moral que implica dar cuenta a los estudiantes de los conocimientos que nos instalan en el lugar del saber. El "jugar al alumno" como parte de una puesta dramática en el momento en que todas las tensiones están puestas en la demostración del saber no se realiza para identificarnos con los alumnos en la tensión, sino para dar cuenta, en situación de prueba, de nuestros conocimientos. Por eso la propuesta, más que identificarse con el estudiante, se organiza para distanciarse de él. El docente, en su apuesta a "más del 10", en la rapidez con que elabora el problema a resolver, juega a reconocer en su producción el lugar de experto como categoría moral que le permita seguir posicionándose en el lugar del saber, pero a la vista de sus alumnos y, por tanto, desde la transparencia de su conocimiento.

Entendemos además que las resoluciones de los expertos proveen de una base para que el alumno analice comparativamente sus producciones, sus aciertos, sus logros y se explique sus dificultades. En ningún caso sostenemos que las evaluaciones de los docentes debieran ser las que expresan la más alta calificación. No podemos comparar lo incomparable. Seguramente, el análisis experto, si el examen es un buen examen, será diferente al del alumno. En los casos en que el examen sólo proponga la explicitación de conocimientos almacenados es probable que no podamos, en algunos de ellos, diferenciar expertos de novatos.

En el caso planteado, los alumnos pudieron reconocer también cómo la formulación condicionó la respuesta. Un efecto no previsto fue que generaron propuestas poco complejas para no perjudicar a los compañeros. La sencillez de la propuesta no les permitió, en muchos casos, reconocer los conocimientos que se habían adquirido, expresar el pensamiento en toda su complejidad o realizar un

análisis creativo. Esto se pudo reconocer comparándolo con el examen que propuso el docente que favoreció estas expresiones del pensamiento complejo. "Lo ético" mal entendido tuvo su expresión en la propuesta. No perjudicar en la evaluación fue determinante de algunos exámenes que no pudieron dar cuenta de todos los saberes construidos.

Las tres últimas etapas, más esta última consideración, les permitieron a los alumnos adquirir nuevas explicaciones respecto de sus conocimientos. Fueron construyendo una calificación y reconstruyendo razones frente al reconocimiento de sus saberes y de algunas dificultades que tuvieron en la concreción de la propuesta.

Entendemos que lo moral, como categoría de análisis, adquiere su mejor expresión en una situación práctica, da cuenta de cómo un docente permite que cada alumno asuma la responsabilidad de construir su evaluación entendiendo que esta propuesta permite superar la clásica condición del alumno frente a la evaluación que implica: dejarse evaluar por otros. La participación activa permite a los estudiantes asumir la responsabilidad de la evaluación. Sin embargo, no se lo deja librado a su propio albedrío ya que se lo ayuda a construir un camino responsablemente y a favorecer la adquisición de información relevante respecto de las realizaciones personales. Implica también no adjudicarse por parte del docente el "derecho" a la evaluación sino, en un clima de mutuo respeto, ayudar a la construcción del conocimiento evaluativo. Resolver posibles discrepancias implica para el docente reconocer las razones que el alumno utiliza para la evaluación. En todos los casos es condición moral acordar criterios, entender razones y, en especial, comprenderlo como acto de construcción de conocimiento, pero esta vez como una construcción de los docentes, entendiendo las razones del alumno, desde el propio corazón del reconocimiento de que él también está implicado en lo que el alumno aprendió. Aprenden los alumnos respecto de lo que aprendieron, y aprenden los docentes aquello que los alumnos aprendieron en un acto de mutua implicación, donde la condición moral del aprendizaje nos reconoce en este espejo de mutuos reflejos y por el que, paradójicamente, se recupera un lugar de par. Si en la evaluación se invertían las relaciones de saber en poder, éste es el momento de recuperar los saberes y otorgarle poder al estudiante en un acto que recupera el

lugar de la justicia al reconocer que quien mejor sabe cuánto sabe es el mismo que aprendió.

La evaluación en la buena enseñanza

A lo largo de este capítulo hemos reconocido situaciones paradójicas, contradicciones o controversias en relación con la evaluación de los aprendizajes. También nuevas consideraciones teóricas que permiten reconocer perspectivas diferentes de la evaluación de los aprendizajes, tales como el análisis que nos permitieron efectuar las derivaciones de los postulados de Jerome Bruner.

Howard Gardner (1995) señala también las características de un nuevo enfoque de la evaluación. En primer lugar, sostiene que debiera ponerse el énfasis en la evaluación y no en el examen, en tanto la primera privilegia la obtención de información en los ámbitos más informales y la segunda debate respecto de los mejores instrumentos para ámbitos neutros o descontextualizados. Esto implica también reconocer la importancia de evaluar simple y naturalmente en los momentos adecuados, utilizando múltiples medidas que favorezcan la expresión de diferentes modos de representación y con materiales interesantes, motivadores y sensibles a las posibles diferencias individuales.

Nosotros reconocemos que el principal desafío a la hora de pensar en la evaluación consiste en construir criterios que nos permitan obtener información válida y confiable. Al estudio de la validez, confiabilidad y practicidad, entre otras cuestiones, se refiere en extenso, en este mismo libro, Alicia Camilloni. Muchas veces los docentes consideramos pertinente evaluar teniendo en cuenta el progreso de cada alumno y por eso dos exámenes iguales obtienen calificaciones diferentes, simplemente porque reconocemos esfuerzos diferentes. Se constituyó en criterio, el esfuerzo o el logro del estudiante. En esos casos, simplemente, se ponderó el resultado del examen por el reconocimiento personal del estudiante. En otros casos se adopta un criterio normativo que nos hace evaluar según lo que el grupo responde. El mejor trabajo obtendría la nota máxima y desde esa construcción comparativa evaluamos al resto del grupo. Implica analizar y ordenar de mejor a peor todos los exámenes y ante

una duda releer los próximos para volver a confeccionar el orden. Desde esta perspectiva, la enseñanza obraría siempre independientemente de las condiciones naturales de los estudiantes, de sus experiencias y sus logros. Lo que se somete a comparación es la producción con independencia de cualquier otro reconocimiento.

La construcción de criterios, el análisis del progreso de los estudiantes o la construcción de la calificación desde el rendimiento máximo en un grupo, constituyen las maneras habituales de calificar. El problema reside en que muchas veces utilizamos una u otra forma para justificar la nota que hemos construido previamente al examen del alumno. Es así como a uno lo evaluamos por el progreso, a otro porque es el mejor examen y a un tercero por la creatividad de su propuesta. Con esto queremos expresar que no negamos la posibilidad de incorporar criterios diferentes de manera coherente, sino que rechazamos toda forma de evaluación que implique la búsqueda confirmatoria de nuestras suposiciones.

Entendemos que una buena evaluación requiere la formulación y explicitación de antemano de los criterios que se utilizarán para dar cuenta del nivel de la producción: reproducción de información –obtenida en clases o a partir de lecturas–, resolución original de problemas o resolución que da cuenta de un sistema de aplicación, creatividad u originalidad en la respuesta, reconocimiento de niveles diferentes de análisis en lo que respecta a la profundización temática, etcétera. Construir y exponer los criterios también permiten reconocer si existen criterios implícitos referidos a las normas de presentación, prolijidad u ortografía, y cuál es el valor de estos criterios. Y si, además, se reconoce el particular crecimiento de los estudiantes.

En el caso planteado en las páginas anteriores, los criterios que utilizó la cátedra no se explicitaron y fue parte de la situación de aprendizaje que los estudiantes analizaran la resolución de los casos con los criterios que posibilitaron el análisis de los docentes. Eran los criterios personales los que permitían un análisis diferenciador y una primera evaluación. Al leer las correcciones a su trabajo se volvía a recuperar el análisis de las posibilidades individuales, pero siempre contrastado frente a lo que significa una buena resolución. Es el propio alumno el que pondera su progreso al construir su evaluación diferenciando producciones y reconociendo las dificul-

lades por las que atravesó para la concreción de la propuesta. De esta manera combinamos criterios y puntos de vista, pero no para reproducir la nota anticipada, sino para entender toda la complejidad que implica el acto de evaluar. Transformamos la evaluación en un acto de construcción de conocimiento en el que no comparamos lo incomparable, sino que cada alumno, a partir de conocer la buena resolución –la del experto–, construyó su propio conocimiento referido a su actuación, contemplando sus hallazgos y sus dificultades. En otras oportunidades, compartir con los estudiantes la construcción de criterios evaluativos para las evaluaciones, reconociendo en todos los casos su valor frente al conocimiento, genera un auténtico desafío a la hora de evaluar y recupera la evaluación como un nuevo lugar posibilitador de la buena enseñanza.

Bibliografía

Baudelot, Christian y Establet, Roger: *El nivel educativo sube*, Madrid, Morata, 1990.

Bruner, Jerome: *La educación, puerta de la cultura*, Madrid, Visor, 1997.

De Alba, Alicia: *Evaluación curricular. Conformación conceptual del campo*, México, UNAM, 1991.

Díaz Barriga, Ángel: *El examen: textos para su historia y debate*, México, UNAM, 1993.

Eisner, E.: "Cognition and Representation", *Phi Delta Kappa*, vol. 78, nº 5, enero de 1997.

Fernández Pérez, Miguel: *Evaluación y cambio educativo: el fracaso escolar*, Madrid, Morata, 1986.

Gardner, H.: *Inteligencias múltiples*, Barcelona, Paidós, 1995.

House, Ernesto R.: *Evaluación, ética y poder*, Madrid, Morata, 1994.

Linaza, J.: "Jerome S. Bruner. La pasión por el conocimiento", *Cuadernos de pedagogía*, nº 243, enero de 1996.

Lipman, Matthew: *Pensamiento complejo y educación*, Madrid, Ediciones de la Torre, 1997.

Litwin, E.: *Las configuraciones didácticas. Una nueva agenda para la enseñanza superior*, Buenos Aires, Paidós, 1997.

———: "La evaluación como una explicación ecológica de la actividad en el aula", *Novedades Educativas*, Buenos Aires, año 7, nº 57.

McCormick, Robert y James, Mary: *Evaluación del curriculum en los centros escolares*, Madrid, Morata, 1996.

Mercer, N.: *La construcción guiada del conocimiento*, Barcelona, Paidós, 1997.

Perkins, David: *La escuela inteligente*, Barcelona, Gedisa,

Perrenoud, Philippe: *La construcción del éxito y del fracaso escolar*, Madrid, Morata, 1990.

Resnick L. y Klopfer, L.: *Currículum y cognición*, Buenos Aires, Aique, 1996.

Rosenthal R. y Jacobson L.: *Pigmalión en la escuela*, Madrid, Marova, 1980.

Santos Guerra, Miguel A.: *Hacer visible lo cotidiano*, Madrid, Akal, 1990.

——: *La evaluación: un proceso de diálogo, comprensión y mejora*, Málaga, Ediciones Aljibe, 2ª ed., 1995.

¿ES POSIBLE MEJORAR LA EVALUACIÓN Y TRANSFORMARLA EN HERRAMIENTA DE CONOCIMIENTO?

Susana Celman

A Mario Garo, con quien inicié mis preguntas sobre este tema. A Gloria, Milagros y Marta, con quienes hoy, en la cátedra, comparto el intento de construir algunas respuestas.

Introducción

Escribir sobre el tema de la evaluación educativa suele ser una invitación a ceder a, por lo menos, una de las tres siguientes tendencias. La primera, caer en la tentación de desarrollar un discurso a la vez complejo y abstracto acerca de sus orígenes, trayectoria y connotaciones actuales. Esto puede resultar interesante y aun importante para suscitar reflexiones sobre la construcción histórica del concepto y sus relaciones dentro del campo educativo y fuera de él. Sin embargo, es posible que un enfoque de esta naturaleza importe más para los estudiosos de la pedagogía que para los docentes que trabajan en nuestras escuelas, que son los destinatarios principales del contenido de estas ideas y reflexiones.

El otro peligro es el opuesto. Consiste en reducir el foco de atención solamente al análisis, construcción y elaboración de propuestas concretas destinadas a mostrar y ejemplificar, en el campo de las prácticas áulicas, una serie de metodologías e instrumentos. Un escrito de estas características se realiza, generalmente, con la pretensión de haber elaborado una propuesta innovadora y carente de dificultades interpretativas, para que pueda ser utilizado por los profesores de distintas asignaturas y niveles. Ubicarnos en esta opción significaría adoptar, aun sin quererlo, una concepción tecnicista dentro del campo pedagógico. Se reduce así el trabajo docente al seguimiento y aplicación de técnicas prescritas por los "especialistas". Se obvian o no se desarrollan los principios teóricos

que las sustentan, a partir de los cuales los profesores podrían estar en condiciones no sólo de elegir, sino de generar nuevas alternativas dentro del enfoque más amplio de educación y de evaluación en particular por el que hayan optado.

Una tercera tendencia, bastante frecuente y que juzgo con serias posibilidades de banalización y superficialización, es intentar responder a las preguntas *¿qué?, ¿cuándo?, ¿cómo?* evaluar, de manera directa y específica. Ésta, al igual que la anterior, suele quedarse en una prolija y ordenada descripción de cierto número de cuestiones, que se presentan bajo el formato de aparentes respuestas a preguntas, también aparentemente sustantivas. Dentro de esta tendencia, suelen encontrarse textos que aíslan la problemática evaluativa, enfocándola como un sencillo acto, para cuya "solución" sólo hace falta precisar algunas cuestiones que facilitarán, luego, actuar con precisión y efectividad.

La intención de este capítulo es plantear el tema desde otro lugar. A tal fin, se ha elegido a modo de encuadre general del trabajo, la presentación de una serie de criterios y principios, que poseen un cierto grado de generalidad y abstracción, analizándolos en un breve desarrollo teórico que intenta explicitarlos y, a la vez, acotarlos al tema propuesto. Asimismo, para cada uno de ellos se plantearán ejemplos pertinentes al campo de la evaluación educativa, posibles de ser relacionados con situaciones áulicas e institucionales. Nuestro objetivo principal es, entonces, poder generar y responder a los intereses y las necesidades de docentes preocupados por la calidad educativa de su trabajo, preocupación que comprende, entre otras, a las prácticas evaluativas.

La hipótesis con la que nos vamos a manejar es que, efectivamente, es posible transformar a la evaluación en una herramienta de conocimiento, en especial para los profesores y para los alumnos, si es que se toman en consideración algunas cuestiones y se preservan y desarrollan otras.

ALGUNOS PRINCIPIOS COMO GUÍA DE REFLEXIÓN

> a) La Evaluación no es ni puede ser un apéndice de la enseñanza ni del aprendizaje; es parte de la enseñanza y del aprendizaje. En la medida en que un sujeto aprende, simultáneamente evalúa, discrimina, valora, critica, opina, razona, fundamenta, decide, enjuicia, opta... entre lo que considera que tiene un valor en sí y aquello que carece de él. Esta actitud evaluadora, que se aprende, es parte del proceso educativo que, como tal, es continuamente formativo (Álvarez Méndez, 1996).

Con este principio se pretende sacar a la evaluación del lugar en el que comúnmente se la ubica: un acto final desprendido de las acciones propias de la enseñanza y el aprendizaje. Se opone a adjudicar a la evaluación el papel de comprobación, de constatación, de verificación de unos objetivos y unos contenidos que deben, por medio de pruebas o exámenes, ser sometidos a un acto de control que permita establecer el grado en que los alumnos los han incorporado.

Por el contrario, lo que aquí se pretende destacar es que las actividades evaluativas –en un contexto de educación en el que se respete la constitución del sujeto de la enseñanza y del aprendizaje como sujetos capaces de decisiones fundadas– se constituyen y entrelazan en el interior mismo del proceso total.

> La función educativa de la escuela requiere autonomía e independencia intelectual, y se caracteriza precisamente por el análisis crítico de los mismo procesos e influjos socializadores, incluso legitimados democráticamente. La tarea educativa de la escuela se propone, por tanto, la utilización del conocimiento y la experiencia más depurados y ricos de la comunidad humana para favorecer el desarrollo consciente y autónomo de los individuos y grupos que forman las nuevas generaciones de modos propios de pensar, sentir y actuar. En definitiva, la potenciación del sujeto (Pérez Gómez, 1997).

La condición para que esto ocurra es que se conciba la tarea educativa como una propuesta que se pone a consideración de sus

actores, quienes la ejercen con autonomía responsable y transformadora. Por ello, es posible entenderla menos como un tranquilo y organizado campo de certezas, y mucho más como un apasionante espacio generador de interrogantes.

Un enfoque prescriptivo de la educación supone que el docente limita su trabajo a ejecutar en el aula las indicaciones que otros han confeccionado para él. Esta suposición no sólo es éticamente insostenible sino, también, empíricamente falsa. Quienes tengan alguna experiencia en la profesión docente, saben que si hay algo que caracteriza a esta tarea es, precisamente, su imposibilidad de ser realizada siguiendo pautas muy específicas y analíticamente prescritas. Todo currículo y cualquier metodología, por organizados que estén, son sólo propuestas y sugerencias que se transforman por la acción mediadora de las instituciones y sus docentes.

Analizar, criticar, juzgar, optar, tomar decisiones no es algo ajeno a la cotidianidad de profesores y alumnos. ES el núcleo mismo de este trabajo con el conocimiento y ES una actividad evaluativa. De otra forma, creemos que podrá haber instrucción, en el sentido que marca su etimología (*"informar"*), pero no un verdadero acto de conocimiento.

Veamos en un ejemplo esto que acabamos de afirmar teóricamente.

Supongamos que nos ubicamos en temas relacionados con la astronomía y, dentro de ella, específicamente con el sistema solar. Desde un inicio, como docentes, nos surgen distintas posibilidades:

- ¿Comenzaremos por las nociones referidas a la totalidad, al sistema, o, por el contrario, iniciaremos el proceso de enseñanza a partir de la noción de "planeta", "satélite", etcétera?
- ¿Qué "puerta" del conocimiento será conveniente utilizar para favorecer un aprendizaje más significativo: ¿la verbal?, ¿la gráfica?, ¿la informática?
- ¿Qué camino seguir y en qué orden?: ¿primero el tratamiento de nociones astronómicas y luego históricas, o al revés?
- ¿De qué modo los alumnos podrán comprender el concepto de "modelo" como modo de pensamiento y método científico?
- ¿Con qué nociones matemáticas relacionar este tema? ¿Con escalas? ¿Con el cálculo de velocidad, tiempo, etcétera? ¿Y físi-

cas? La ley de gravedad, por ejemplo. También éticas: ¿el uso de animales en experimentos, la contaminación de la atmósfera?

- ¿Les pediremos a los alumnos, como "prueba de su aprendizaje", que reproduzcan el esquema del sistema solar, con los nueve planetas, adjudicándole el nombre a cada uno de ellos, o quizás es posible pensar en la construcción de un modelo a escala? ¿Qué conocimientos y saberes se ponen en juego en uno y otro caso?...

Las decisiones de los profesores, previamente razonadas o tomadas al correr de los acontecimientos, implican analizar, criticar alternativas, juzgar sobre la base de ciertos criterios y, por último, optar. Pero paralelamente, los alumnos también realizan estas acciones evaluativas. Ellos también analizan, critican, discuten, discriminan, juzgan... quizás a partir de interrogantes y con criterios y finalidades no del todo coincidentes con los propósitos del docente, pero lo hacen.

Y por último, este principio también advierte que este modo de proceder con el conocimiento se aprende, es decir, puede verse favorecido y estimulado por procesos intencionales y sistemáticos encarados en función de un objetivo educativo explícito ("desarrollar en los sujetos actitudes evaluativas y críticas respecto de los contenidos del aprendizaje") pero también son el producto de convivir en un ambiente educativo que se caracterice por este estilo de trabajo intelectual.

El mejor método que un profesor puede utilizar para que sus estudiantes desarrollen formas activas y creativas de aprendizaje es transparentar, en sus clases, los procesos que él mismo puso en juego al aprender: sus dudas, sus criterios, sus opciones, sus hipótesis. Es posible comentarles: "Yo dudaba si comenzar por este o aquel tema... Creo que este enfoque es mejor porque... Elegí este modo de trabajarlo basándome en... Veremos si con esta actividad se logra... Ésta es una de las maneras de entender el tema; puede haber otras... Estemos atentos a los fundamentos...", etcétera.

De este modo, los alumnos aprenden epistemología, democratización de las relaciones interpersonales, actitudes no dogmáticas hacia el conocimiento... Enseñanza y aprendizaje de los procesos de evaluación característicos de la relación de un sujeto crítico con el conocimiento.

b) La mejora de los exámenes comienza mucho antes, cuando me pregunto: "¿Qué enseño? ¿Por qué enseño eso y no otras cosas? ¿De qué modo lo enseño? ¿Pueden aprenderlo mis alumnos? ¿Qué hago para contribuir a un aprendizaje significativo? ¿Qué sentido tiene ese aprendizaje? ¿Qué otras cosas dejan de aprender? ¿Por qué?".

Este segundo principio está en íntima conexión con el primero. Surge, en gran medida, de las propias experiencias concretas de nuestro trabajo en Talleres con docentes, organizados en torno del tema de la evaluación educativa.

La demanda para la realización de dichos Talleres, desde las escuelas, está fundamentada en necesidades que suelen expresarse con frases tales como: "Queremos mejorar nuestra forma de evaluar", "No sabemos si evaluamos correctamente", "A nosotros no nos han enseñado estos temas de la evaluación", "En esta escuela no nos ponemos de acuerdo; cada uno evalúa como quiere", "Es necesario consensuar criterios comunes"...

Uno de los primeros puntos que ponemos a consideración de los participantes en dichos Talleres es que los métodos y las técnicas de evaluación mejoran su calidad educativa –es decir, su potencial educador– cuando forman parte de un proceso más amplio y más complejo que, a su vez, ha mejorado.

En otras palabras: se intenta mostrar que las pruebas de evaluación de los aprendizajes serán valiosas, en primer lugar, en tanto nos permitan conocer la manera y el grado de apropiación que los estudiantes han realizado de un conocimiento, que se considera importante y digno de ser conocido. Por el contrario, de nada valdrían sofisticados sistemas de evaluación aplicados a dar cuenta de contenidos poco significativos y superficiales.

Un ejemplo de lo dicho sería preocuparse y destinar un tiempo considerable de trabajo docente a la preparación de una compleja prueba de evaluación destinada a conocer el grado de información que poseen los alumnos acerca de los nombres y las fechas que indican el principio y fin de las distintas "Edades" históricas, en lugar de analizar, antes, qué es lo importante, desde un punto de

vista educativo, respecto de esta periodización que no deja de ser, después de todo, uno de los modos posibles de comprender la Historia.

Se pretende fundamentar cómo la escuela y los docentes, en forma individual y/o grupal, tienen un espacio de decisión sobre los contenidos de la enseñanza. Los currículos, normalmente, contienen "contenidos mínimos" cuyo desarrollo más o menos extenso, más o menos complejo, más o menos interconectado, depende de la decisión del grupo educativo formado por profesores y estudiantes.

Por supuesto, esto trae aparejado el problema de tener que optar, y será conveniente hacerlo a partir de ciertos criterios que den cuenta de las razones que motivaron tales decisiones (por ejemplo, habrá que decidir acerca de lo que quedará incluido dentro de los contenidos a enseñar y a aprender y, por ende, los que no serán seleccionados).

Pero, no sólo serán valiosos los exámenes que pretendan evaluar ciertas temáticas, a su vez potencialmente valiosas, sino que tal cualidad depende también del tipo de conocimiento que hayan promovido, y de la calidad del sistema de evaluación para ponerlo de manifiesto. Es posible pensar en construir instancias evaluativas capaces de evidenciar los procesos de sistematización de información, indagación, problematización, relaciones de categorización, generalización, diferenciación, inducción y deducción de principios, aplicación y creación de procedimientos, resolución de problemas, etcétera, si se ha trabajado en tal dirección antes, durante el periodo de enseñanza y aprendizaje.

Pero, al mismo tiempo, al diseñar las actividades específicamente destinadas a la evaluación de modo tal que los estudiantes pongan en juego estos procesos cognitivos, se propiciará que se generen nuevos aprendizajes, como resultado de las nuevas relaciones desencadenadas por esta situación. Es decir, según el tipo de cuestiones que se les plantean a los alumnos durante una prueba, por ejemplo, éstos pueden verse llevados a crear otros "puentes cognitivos" (Ausubel, Novak y Hanesian, 1986) para resolver esta situación, distintos de los que había elaborado durante el período de enseñanza y aprendizaje.

Será factible entender, entonces, que los así llamados "contenidos procedimentales" son una dimensión del "material" con el que

trabajamos en el aula, un modo de enfocar aspectos o cuestiones que sólo artificialmente pueden disociarse del tratamiento que se realice con el objeto de conocimiento, en tanto estos procederes intelectuales lo constituyen como tal. En efecto, al menos una dimensión de dichos contenidos se relaciona con el modo en que se aprende y el modo en que se enseña, es decir, el tipo y calidad de las "actividades mentales" que ponen en juego alumnos y profesores para conocer.

Llevemos esto a una situación concreta para facilitar su comprensión. Para ello recurriremos a la Historia como asignatura perteneciente a las Ciencias Sociales, en un nivel del sistema educativo correspondiente a los primeros años de la enseñanza media.

Hay un acuerdo generalizado en esta disciplina acerca de que centrar el trabajo pedagógico en un enfoque fáctico, que tome como eje prioritario la identificación de "hechos", será poco significativo para los estudiantes, en tanto no sólo reduce sus posibles relaciones, derivaciones y transferencias sino que, por eso mismo, se olvida más rápidamente.

Además, los alumnos presentan durante sus actividades de conocimiento, en general, dificultades propias de su nivel de desarrollo evolutivo; y en particular, dificultades para el aprendizaje de la Historia. Dichas dificultades se expresan, por ejemplo, en problemas relacionados con el manejo del tiempo y el espacio, como nociones constitutivas de los procesos históricos. Por otra parte, a esto se le suman las provenientes de los esfuerzos de descentración necesarios para la comprensión de otras culturas, distantes geográfica y temporalmente, de la que los alumnos están viviendo en la actualidad. Con relativa frecuencia, se suele intentar comprender y juzgar acontecimientos y procesos propios de otras comunidades sociales y que acontecieron en otras épocas, con las ideas, conocimientos, valores y perspectivas inherentes al aquí y ahora.

Sin embargo, es posible y necesario trabajar con los alumnos los conceptos básicos que hacen a esta disciplina, sin los cuales será imposible librarnos de la tendencia a la simplificación esquemática y la excesiva abstracción.

Para ello, se podrá diseñar una serie de actividades que:

- planteen problemas que requieran el desarrollo de conocimientos y habilidades;

- sean susceptibles de tratamientos diversos y distintos niveles de resolución;
- permitan su expresión a través de formas alternativas;
- exijan el manejo de información precisa y rigurosa,

y a la vez,

- faciliten la apertura interpretativa;
- soliciten la consulta a distintas fuentes de información y requieran el ordenamiento y sistematización de los datos;
- permitan la elaboración de redes conceptuales;
- promuevan la autoevaluación y la coevaluación grupal y de la tarea; etcétera.

Los aspectos antes señalados se constituyen, a la vez, en los criterios de evaluación del trabajo de los alumnos. Cada uno de ellos forma parte de un conjunto de objetivos valiosos propios del aprendizaje de las Ciencias Sociales y orientan la fundamentación de opciones metodológicas de su enseñanza y su aprendizaje.

c) No existen formas de evaluación que sean absolutamente mejores que otras. Su calidad depende del grado de pertinencia al objeto evaluado, a los sujetos involucrados y a la situación en la que se ubiquen.

En algunas épocas de la historia de la evaluación educativa se creyó que existían formas de evaluación que eran indudablemente superiores a otras. Es decir, se ponía el acento en la manera de construir, de confeccionar el/los instrumentos con los cuales se procedería a evaluar, poniendo especial énfasis en ciertos atributos que debían reunir dichos instrumentos. Con ello se pretendió diferenciar la evaluación "científica" de aquella más "ingenua" o "intuitiva". La primera suponía ciertos saberes técnicos y la segunda era la practicada por los docentes en general, basándose en su experiencia de trabajo en el aula con los diferentes grupos de alumnos.

Esto ocurrió especialmente con el caso de las así llamadas "pruebas objetivas", las cuales recibían el carácter de "cientificidad" a partir de su semejanza con los principios de la psicometría y los tests. Toda una etapa de la historia de la evaluación estuvo marcada por esta problemática, que se extiende a partir de fines de la década de los 40 en Estados Unidos y Europa, y prácticamente dos décadas más tarde en nuestro país (véanse Stufflebeam y Shinkfiek, 1987; Angulo Rasco, 1990, entre otros).

En efecto, una de sus cualidades, la supuesta "objetividad", se esgrimía como prueba de su indudable calidad respecto a otras maneras de evaluar que se mostraban como "subjetivas" y, por ende, poco rigurosas y mucho menos científicas. Entre otras cosas, esto contribuyó a consolidar el prejuicio referido a la falta de capacidad y capacitación de los profesores para la tarea evaluativa y a derivarla, implícita o explícitamente, hacia otros sujetos ubicados en otros lugares del sistema educativo.

Sin embargo, el análisis crítico de estas pruebas mostró que la "objetividad" se restringía al momento de la corrección de los ítemes, dado que su construcción prevé sólo marcar la alternativa correcta o seleccionar una entre varias –para nombrar las formas más usadas–, de modo tal que cualquier persona, y hasta una máquina que contenga las respuestas esperadas, podría corregirlas sin dudar, es decir, marcar y contar cuántos aciertos tuvo cada alumno. Pero se advertía que dicha "objetividad" es inexistente en el momento de construir la prueba, de decidir qué cuestiones abarcaría, cuáles quedarían fuera, qué peso –valor o puntaje– tendría cada aspecto, etcétera.

Además, se señalaron importantes problemas que se derivaban de su uso:

- la fragmentación excesiva de los contenidos, al dividir un tema o cuestión en un gran número de ítemes para resolver;
- las limitaciones en el caso de las Ciencias Sociales y Humanas, donde los aspectos más significativos de estas disciplinas raramente pueden plantearse en formulaciones cerradas y en opciones excluyentes entre sí;
- su centralización preponderante en evaluar los resultados, la respuesta final;

- las dificultades en el análisis de los caminos por los cuales llegó el alumno a elegir una de las opciones que se le presentaron, etcétera.

Estas y otras críticas se encuentran ampliamente difundidas en la bibliografía especializada. Además, sólo nos interesaba este caso a modo de ejemplo, para presentar una de las polémicas típicas del área de la evaluación y algunos argumentos que demuestren la falta de unanimidad respecto de la superioridad de estos métodos sobre otros.

Queremos decir con esto que la respuesta más consistente que estamos en condiciones de dar frente a la pregunta "¿Qué método de evaluación es mejor?" es "Depende en qué caso" y "Depende para qué".

Nuestra postura es que los objetos de evaluación son construidos gracias a las preguntas que les formulemos y las finalidades que se les hayan atribuido. Por lo tanto, variarán de acuerdo con ellas. "Lo que se evalúa" no son "cosas" con existencia e identidad independiente de quienes las valoran.

Es muy diferente una evaluación destinada a comprobar qué se ha retenido de un tema que se estudió consultando un texto único, que aquella otra destinada a conocer el tipo de relaciones que el estudiante ha sido capaz de hacer entre distintos autores, al interior de dicho tema y con otros, las opiniones que le merecen, las aplicaciones a situaciones diferentes, las preguntas en las que se quedó pensando, etcétera. Adviértase que, en este caso, estamos refiriéndonos al mismo tema, pero al trabajarlo de diferentes maneras también deberán ser distintas las formas con las que se lo intente evaluar, porque son otras las cuestiones que se desea conocer.

Intentaremos mostrar esto con dos ejemplos: uno, referido a un tema de una asignatura, y el otro, relacionado con el desarrollo de ciertas habilidades.

Tomemos por caso que nuestra intención es evaluar el conocimiento que los alumnos han logrado respecto a un tema como "El azúcar". A primera vista parecería que no hay dificultad en ubicar el objeto al cual vamos a referir la evaluación: el azúcar.

Pero inmediatamente nos damos cuenta de que esto no es tan sencillo. Nos preguntamos: ¿qué hemos trabajado en relación con

este tema? ¿Lo enfocamos desde las Ciencias Biológicas prestando atención a la conformación de la caña, sus características, su clasificación botánica? Y en este caso, ¿el foco estuvo puesto en los elementos morfológicos constitutivos o en los procesos fisiológicos?

Las formas de evaluar estos aprendizajes deberán variar de acuerdo con las áreas conceptuales trabajadas y las operaciones cognitivas priorizadas. En el primer caso del ejemplo anterior será pertinente indagar sobre la precisión de la información que posean los estudiantes respecto de los elementos constitutivos y su denominación correcta. Se podrá optar por formas gráficas, por preguntas dirigidas a captar datos simples como ¿qué?, ¿dónde?, etcétera. En el segundo, los interrogantes se dirigirán mucho más a poner de manifiesto causas, razones, situaciones problemáticas, y podrían ser evaluados en el transcurso de experiencias concretas.

¿Y si en lugar de posicionarnos en la Biología lo hicimos en la Química y los procesos de composición y transformación de la sustancia? Para su evaluación se deberán diseñar operaciones y problemas distintos de los anteriores, que deben dar cuenta de su especificidad.

¿Y si lo estudiamos desde la Sociología, la Economía y la Política? Seguramente cambiarán las fuentes de información a las que se debe recurrir (periódicos, estadísticas, testimonios orales, películas) y serán otros los procesos de conocimiento puestos en juego. Cambiarán también los modos de evaluación, puesto que habrá que elegir aquellos que faciliten conocer estos procesos: el coloquio, la redacción de informes, la producción de material gráfico o audiovisual, etcétera.

Pasando al otro ejemplo, supongamos que se trata de trabajar, en una escuela rural, con contenidos del área de conocimientos tecnológicos relacionados con la producción de lácteos. Allí, es posible que juzguemos importante enfatizar la relación teoría-práctica. Trabajaríamos sobre los conocimientos transmitidos empíricamente por el medio familiar y cultural de los alumnos. Es decir, recuperaríamos lo que los estudiantes ya saben por haberlo aprendido en sus casas, trabajando con los adultos de su entorno familiar para, desde allí, iniciar el proceso de engarzar conceptos, principios y procedimientos sustentados científicamente. La otra cuestión será llevarlos al plano de las realizaciones técnicas de

producción y volver a recuperar, desde ese lugar, las fundamentaciones teóricas.

En este caso, si bien es posible pensar en una evaluación de los aprendizajes mediante una prueba de lápiz-papel, ésta sólo nos brindará información acerca de los saberes teóricos y verbales que hayan desarrollado los estudiantes, será muy pobre para juzgar el "saber hacer" tecnológico. Para esto, una evaluación pertinente deberá acercarse al plano de las realizaciones concretas, mediante la observación, el registro, la entrevista focalizada, la resolución de problemas prácticos y el diálogo explicativo posterior.

En síntesis: esta relación entre el enfoque con el cual nos dirigimos a un área de conocimiento, las preguntas que le formulemos, las operaciones cognitivas que se potencian, son elementos que marcan y determinan en gran medida el tipo y forma que debe adoptar la evaluación, para que se constituya en una evaluación de calidad.

d) Si el docente logra centrar más su atención en tratar de comprender qué y cómo están aprendiendo sus alumnos, en lugar de concentrarse en lo que él les enseña, se abre la posibilidad de que la evaluación deje de ser un modo de constatar el grado en que los estudiantes han captado la enseñanza, para pasar a ser una herramienta que permita comprender y aportar a un proceso.

La lectura del interesante libro de Newman, Griffin y Cole, *La zona de construcción del conocimiento* (1991), inspirado en parte en los aportes de Vigotsky pero enriquecido con el trabajo de largas observaciones en las aulas de escuelas de Estados Unidos, nos ha inducido a la formulación de este principio. En el capítulo V los autores plantean tanto la necesidad cuanto las posibilidades de este cambio de enfoque. A partir de los desarrollos teóricos y las reflexiones críticas que contienen los capítulos anteriores, se muestra en éste el tipo de intervención que los docentes pueden adoptar para seguir la pista de los cambios cognitivos que van ocurriendo en los alumnos durante las clases.

Se trata de lo que los autores denominan "evaluación dinámica" o "evaluación a través de la enseñanza". A diferencia de los procedimientos destinados a medir la realización satisfactoria de una tarea en un momento determinado, estos autores sugieren:

> En vez de proponer una tarea a los niños y medir hasta qué punto la hacen mejor o peor, podemos proponérsela y observar cuánta ayuda y de qué tipo necesitan para terminarla satisfactoriamente. De este modo no se evalúa al niño en forma aislada. Se evalúa el sistema social formado por el profesor y el niño para determinar cuánto ha progresado.

Una de las tareas del docente, en este enfoque, es determinar –evaluar– cuándo es conveniente, necesario y posible intervenir para promover el cambio cognitivo.

Este tipo de propuesta involucra la presentación de situaciones problemáticas que el estudiante debe resolver trabajando con determinados materiales. El docente observa las acciones que ellos realizan por sí solos. Cuando un alumno llega a un punto que no puede continuar, el profesor, mediante interrogantes o indicaciones, le da pistas acerca de por dónde puede seguir para continuar haciendo las comprobaciones por su cuenta. No se trata de decirle lo que debe hacer, sino de brindarle alguna ayuda para que pueda continuar desarrollando sus actividades por su cuenta, en forma independiente.

La "evaluación dinámica" comprende dos aspectos: el primero evalúa el estado actual del niño en relación con la zona disponible para la adquisición del concepto. El segundo evalúa la "modificabilidad", la disponibilidad del alumno para aprender.

Por ejemplo, supongamos que estamos trabajando con un curso a partir de una situación problemática existente en la realidad: la instalación de un basurero nuclear en las cercanías de la ciudad. Encarado este tema desde las Ciencias Sociales, inmediatamente surge un gran número de interrogantes: ¿se trata de un problema político, jurídico, ecológico, económico, legal, cultural, técnico o geológico?

Durante los primeros encuentros de los alumnos con esta temá-

tica, y mediando un enfoque didáctico que incorpore los principios vigotskianos, el profesor podrá avanzar en la primera evaluación, es decir, podrá ir conociendo qué conceptos generales y específicos están disponibles en los estudiantes para construir sus aprendizajes; cómo los relacionan con los nuevos materiales; hasta dónde son capaces de continuar el proceso y cuáles son las estrategias cognitivas que utilizan.

Desde allí, y a partir de reconocer las posibilidades y los obstáculos, es el docente quien debe elegir el tipo de aporte que necesitan de su parte. Habiendo detectado cuáles son los conceptos y las herramientas cognitivas con las que los han construido y los están usando, podrá evaluar cuáles son las facilidades y los límites con que están actuando en el grupo, viabilizando y trabando, respectivamente, su aprendizaje autónomo. Esto le posibilitará aportar sugerencias, indicaciones, informaciones, ideas, procedimientos, según los casos, y juzgar el grado de disponibilidad de sus estudiantes para aprender.

En nuestro país se están retomando en la actualidad las ideas provenientes de la teoría de Vigotsky, las cuales tuvieron, décadas atrás, un lugar importante en las Ciencias de la Educación y en la Psicología Educativa. Por lo menos se ha comenzado a incluirlas en los programas de formación docente y en cursos y seminarios destinados a los profesores. No tenemos datos acerca de su empleo como marco referencial en el trabajo cotidiano en las escuelas e institutos. A pesar de ello, consideramos que su riqueza teórica y su potencia para generar propuestas prácticas justifican su inclusión en este capítulo.

> e) Obtener información acerca de lo que se desea evaluar es sólo un aspecto del proceso evaluativo. Su riqueza y, a la vez, su dificultad mayor consiste en las reflexiones, interpretaciones y juicios a que da lugar el trabajo con los datos recogidos.

Frecuentemente se confunde la toma del "dato" con el complejo proceso de la evaluación. La información proveniente de distintos

procedimientos es el material a partir del cual se inicia, realmente, el proceso evaluativo, no la evaluación misma.

En realidad, el diseño y/o elección de la metodología para captar la información supone, previamente, una toma de postura teórico-epistemológica acerca de la concepción de conocimiento y de su modo de construcción. Esto implica analizar, discriminar y juzgar.

Por otra parte, diseñar instrumentos y aplicarlos, optar por alguna metodología en lugar de otras, nos permitirá tener acceso a una cierta cantidad de información en relación con lo que se desea evaluar. Su capacidad para dar cuenta de diferentes tipos de conocimientos, dependerá de la sensibilidad de dichos instrumentos para captar o capturar los datos buscados. El cuidado en su aplicación (elección del momento, características de la situación, clima creado para su realización, etcétera) también incidirá en la validez y confiabilidad de los datos que se obtengan.

Pero lo que se desea destacar en este punto es que, si bien constituyen un aporte importante, no son el núcleo central de la evaluación.

Pensemos esta idea a partir de un ejemplo ajeno al campo de la evaluación.

Para evaluar un terreno, podemos seguir distintos procedimientos: tomar muestras de tierra para su análisis químico; aplicar instrumentos de medición para establecer sus dimensiones; tomar fotografías en diferentes momentos del día; realizar observaciones desde un helicóptero; mantener entrevistas con los vecinos, etcétera. Pero, tanto la interpretación como la valoración subsiguiente en estos datos dependerán no de los datos mismos, sino de lo que pretendamos hacer con ese terreno.

Por cierto, cada dato cambiará de significación e importancia si al terreno lo necesitamos para construir una vivienda para el fin de semana, poner un establecimiento agrícola o una estación de servicio. Lógicamente, también será diferente el juicio acerca de su calidad, y puede suceder que varias personas arriben a conclusiones diversas sobre el valor del terreno, a partir de los mismos datos, pero con concepciones de uso dispares.

Pero además, en todos los casos, quienes deben emitir el juicio valorativo realizan procesos reflexivos de mayor o menor profundidad y extensión, a partir de los datos que manejan. Dicho proceso

consiste no sólo en relacionar unos con otros, sino también preguntarse por las razones y factores que determinan su estado actual, así como las tendencias de cambios futuros.

Del mismo modo, un docente cuenta con múltiples y particulares fuentes de información que le brindan datos acerca de los procesos de aprendizaje de sus alumnos: sus intervenciones en clase, sus preguntas, la manifestación de múltiples actitudes, sus trabajos, sus exámenes, etcétera.

En las escuelas se construyen diferentes dispositivos para lograr obtener y sistematizar el conocimiento que se va acumulando acerca de las peculiaridades de los procesos educativos que allí ocurren. Docentes y alumnos pueden disponer de diversas fuentes de información al respecto: desde las más pobres en significatividad, como son las listas con las notas obtenidas por un curso, a otras indudablemente más ricas en potencialidades para el conocimiento, como los registros de observaciones de actividades áulicas y los trabajos producidos durante el año.

Pero todos éstos son datos que deberán ser procesados y, sobre todo, interpretados. Las concepciones que se tengan acerca del conocimiento, la enseñanza, el aprendizaje, constituyen marcos referenciales epistemológicos y didácticos que, juntamente con criterios ideológico-educativos y consideraciones acerca del contexto en que se desarrolla el proceso de enseñanza y aprendizaje, actúan a modo de parámetros que guían dicha reflexión y orientan las interpretaciones.

Por ejemplo, desde un enfoque constructivista del aprendizaje, no será bien visto un trabajo evaluativo que consista en verificar el grado en que los estudiantes han sido capaces de sintetizar las ideas principales de un autor. Por el contrario, sí se valoraría que se intentara comprender cómo sus alumnos han construido un sistema categorial para aprender los nuevos conocimientos de una asignatura en particular.

No es posible identificar evaluación con medición de conocimientos, ni confundirla con la aplicación de instrumentos tipo tests. Como dice F. Angulo Rasco (1994):

> Es necesario conocer lo que va a ser juzgado, pero el conocimiento es insuficiente. Porque formular un juicio de valor, y no un juicio

numérico, es un proceso cognitivo mucho más complejo que relacionar [...] puntuaciones con calificaciones. Ambos conceptos o acciones se encuentran en planos intrinsecamente distintos. Creer que las puntuaciones (calificaciones) son los únicos elementos necesarios para formular un juicio de valor, supone desconocer u ocultar la participación, normalmente inconsciente, de otros elementos como los valores que sustentamos, nuestras concepciones educativas y docentes. Sin embargo, no se trata de aumentar la información disponible, o no sólo. Es un error creer que la cantidad de información sobre el alumnado constituye una condición suficiente del juicio, cuando en realidad es únicamente una condición necesaria...

Y más adelante, en el mismo texto, hace mención a una cita de la cual es autor juntamente con J. Contreras y A. Santos Guerra:

Formular un juicio debería ser entendido como un proceso social de construcción, articulado en el diálogo, la discusión y la reflexión, entre todos lo que, directa o indirectamente, se encuentran implicados en y con la realidad evaluada.

Ubicándonos en el interior del aula, las "verdaderas evaluaciones" serán aquellas en las que docentes y alumnos, con la información disponible, se dispongan a relacionar datos, intentar formular algunas hipótesis y emitir juicios fundados que permitan comprender lo que ocurre, cómo ocurre y por qué.

Por ejemplo, frente al análisis de ciertos trabajos producidos por los alumnos, el docente puede preguntar y preguntarse por qué son ésos y no otros los resultados; qué factores han incidido; cuáles, entre ellos, han tenido mayor peso; qué grado de coincidencia hay en el grupo sobre dichos determinantes; qué medidas son posibles y necesarias para mejorar tanto los procesos cuanto los productos.

Con esto no se descarta el hecho de calificar esos trabajos lo más responsablemente que sea posible, si así lo establecen las normas vigentes. Se trata de advertir respecto a que las calificaciones en sí, sólo aportan información sobre el lugar que ocupa cada alumno y su rendimiento en una escala numérica o conceptual. También es posible que se constituyan en una señal que indique, para algunos sujetos, que algo no anda bien. Lo que seguramente no le dicen es qué y por qué "no anda bien". Y menos aún, qué debería ser

modificado. Es decir, se transforman en un signo opaco, que no permite avanzar en la elucidación de los procesos y los motivos. No permite, en suma, aprender.

> f) La evaluación se constituye en fuente de conocimiento y lugar de gestación de mejoras educativas si se la organiza en una perspectiva de continuidad. La reflexión sobre las problematizaciones y propuestas iniciales, así como sobre los procesos realizados y los logros alcanzados –previstos o no previstos–, facilita la tarea de descubrir relaciones y fundamentar decisiones.

Con esto no se niegan las evaluaciones puntuales que tengan por objetivo dar cuenta de un determinado estado o situación. Es posible realizarlas y su calidad específica dependerá, entre otras cosas, de aspectos ya señalados en las páginas anteriores. Su principal función será la de constatar las realizaciones de los alumnos respecto a los objetivos pedagógicos previamente planteados.

Pero la construcción de un juicio evaluativo acerca de los procesos de enseñanza y aprendizaje que ocurren en las escuelas requiere reconocer la especificidad del hecho educativo y, dentro de ella, su carácter procesal, dinámico y multideterminado.

M. Scriven (1967) formuló el concepto de "*evaluación formativa*", para referirse a las actividades concebidas para permitir los reajustes necesarios y sucesivos en el desarrollo de un nuevo programa, manual o método de enseñanza. Posteriormente, este concepto se aplicó a los procedimientos utilizados por los docentes a fin de adecuar sus estrategias pedagógicas, de acuerdo con los progresos y dificultades mostradas por sus alumnos.

Linda Allal, de la Universidad de Ginebra, en un artículo publicado por la revista *Infancia y aprendizaje* (1980) recupera el concepto de Scriven, y define lo que, a su entender, son los rasgos de una evaluación formativa desde una perspectiva cognitivista:

> [...] en una evaluación formativa se intenta ante todo comprender el funcionamiento cognitivo del alumno frente a la tarea propuesta. Los

datos de interés prioritarios son los que se refieren a las representaciones que se hace el alumno de la tarea y a las estrategias o procedimientos que utiliza para llegar a un determinado resultado. Los "errores" son objeto de un estudio en particular en la medida en que son reveladores de la naturaleza de las representaciones o de las estrategias elaboradas por el alumno.

Los actos de evaluación aislados y descontextuados que se realizan a modo de "corte vertical" de dicho proceso poco nos dicen sobre las razones por las cuales éste ha ocurrido de ese modo. Quizá su mayor valor consista en posibilitarnos la elaboración de interrogantes, de hipótesis, de preguntas cuyas respuestas deberán buscarse fuera de dichos actos. Nos permiten detectar algunos aspectos que, en ese particular momento, se "muestran" a nuestra observación: algunas cuestiones acerca de las que pueden responder los alumnos; otras que les ofrecen dificultades; una parte de la información que han asimilado –aquella sobre la cual preguntamos, dado que sobre lo que no preguntamos no sabremos si lo maneja o no–; algunos problemas que puede resolver y otros que no; determinados procedimientos que utiliza en esa instancia, de manera más o menos apropiada; etcétera.

La información proporcionada por estas "evaluaciones verticales" actúa como la radiografía en el proceso de diagnóstico médico clínico: permiten reconocer algunos signos esperables ("normales") y detectar, si los hubiera, los que no lo son ("anormales"). Pero si deseamos arrimarnos a la génesis de su desarrollo, por lo menos habrá que repetir las radiografías un cierto número de veces y comparar unas con otras. Las sucesivas placas van a constituir los fundamentos para el juicio médico acerca de ese paciente: "Va mejor", "Anda bien" o "Ha empeorado".

Pero las cosas no son tan sencillas. En este ejemplo será necesario aclarar algunas cuestiones. En primer lugar, el concepto de "salud" y de "enfermedad" a partir de los cuales se estudiaron las radiografías. En segundo lugar, el proceso de elaboración de hipótesis acerca de cuál es la enfermedad que aqueja al paciente, y en tercer término, el marco teórico y la experiencia práctica con que cuenta el profesional para juzgar los elementos que brindan las placas e interpretarlos.

Sin embargo, no termina aquí el diagnóstico evaluativo. Es probable que no basten estas fuentes de información. Seguramente un buen médico realizará entrevistas con el paciente para confeccionar una completa historia clínica; solicitará datos acerca de su desarrollo evolutivo anterior a esta consulta; reunirá y estudiará los informes de otros colegas y pedirá la realización de nuevos estudios... El análisis reflexivo de estos elementos de juicio posibilitará arribar a un diagnóstico más adecuado y preciso, intentar una explicación fundamentada del problema y elaborar una propuesta de tratamiento.

De modo semejante, en las escuelas, para poder interpretar determinadas pruebas de rendimiento también es necesario explicitar qué vamos a considerar "buen aprendizaje"; qué entendemos por "logros educativos"; qué queremos decir cuando decimos "sobresaliente", "bueno" o "regular"; qué queremos significar con "buena enseñanza"...

A partir de estos parámetros es posible realizar un análisis de distintas situaciones, elaborar las primeras hipótesis y, al igual que en el ejemplo anterior, seguramente el docente que aspire a un conocimiento más amplio y profundo intentará establecer relaciones comparativas entre los datos actuales y los anteriores, analizará las posibles determinaciones que han podido incidir y elaborará algunos principios teóricos para explicar, de manera más o menos comprensiva, el proceso en cuestión. Dialogará con los alumnos y los colegas, buscará información en otros documentos...

Por otra parte, es factible pensar en proponer algunos recursos didácticos que faciliten formas de colaboración e intercambio entre los propios alumnos, que les posibilite desempeñar ellos mismos tareas de observación y registro de las actividades de aprendizaje. Esto permitiría, en grupos numerosos, integrar la evaluación a las actividades pedagógicas, sin centrar sólo en el profesor dicha tarea. Actualmente, la disponibilidad de computadoras en las escuelas puede constituir un recurso técnico adecuado a tal efecto, si se crean programas educativos que faciliten a los estudiantes descubrir las características de su modo de trabajo cognitivo, así como reorientarlo ante los eventuales obstáculos.

Desde allí, y recuperando los datos de una historia personal y grupal, estaremos en situación de elaborar evaluaciones que permi-

tan comprender y valorar, al menos en parte, los procesos educativos que nos involucran a nosotros y a nuestros alumnos.

Una última aclaración antes de terminar este punto: el desarrollo que acabamos de hacer remite en términos generales a una concepción de evaluación de proceso. Entendemos que los rasgos más característicos de este tipo de evaluación no consisten en repetir frecuentemente actividades evaluativas. Por el contrario, se manifiestan en la intencionalidad de analizar y comprender el proceso tal como va ocurriendo, deteniéndose especialmente en el estudio del tipo y cualidad de las relaciones que podrían haber actuado como factores determinantes del mismo. Las hipótesis explicativas que se elaboren de este análisis permitirán no sólo entender qué pasó sino, fundamentalmente, constituirlo en una experiencia educativa para su mejora.

g) La evaluación de las estrategias de aprendizaje puestas en juego durante el proceso de construcción de los conocimientos, es un área de alta potencialidad educativa y con amplias posibilidades de incidencia en la transformación de dicho proceso.

El concepto de *"estrategias de aprendizaje" (E. de A.)* pertenece al campo de la psicología cognitiva. Autores como Monereo y otros (1994), Nisbet y Schuckmith (1987) y Pozo (1993), entre otros, han desarrollado interesantes trabajos al respecto. Este concepto se refiere, en términos generales, a los diversos procedimientos que pone en juego un sujeto al aprender y abarca, para algunos de ellos, *"desde el uso de simples técnicas y destrezas, al dominio de estrategias complejas"* (Pozo y Postigo Aragón, 1993).

La importancia que actualmente se atribuye al desarrollo de las habilidades de "aprender a aprender" se deriva no sólo de concepciones teóricas propias del ámbito educativo, sino de las demandas sociales que parecen requerir la formación de individuos capaces de un mayor manejo autónomo de estas herramientas cognitivas.

Vivimos en un mundo en el que la velocidad de los procesos de evolución y transformación de los conocimientos en general, así

como de los desarrollos tecnológicos en particular, demandan de los jóvenes y adultos que no sólo sean capaces de adquirirlos, sino también construir las E. de A. pertinentes a los distintos objetos de conocimiento.

Nisbet y Schuckmith (1987) definen las E. de A. como *"secuencias integradas de procedimientos o actividades que se eligen con el propósito de facilitar la adquisición, el almacenamiento y/o la utilización de información o conocimientos"*.

En esta definición aparecen dos conceptos importantes: "secuencias integradas" y "elegidas con un propósito". Ambos marcan una diferencia significativa entre las E. de A. y otros tipos de procedimientos, en tanto parecen distinguirlas, diferenciándolas, de las acciones aisladas, mecánicas y automáticas, y las refieren a procesos integrados, que pueden llegar a ser construidos, conocidos y utilizados por el sujeto, en función de determinadas demandas de las tareas de aprendizaje.

El hecho de que incluyan ciertas técnicas o destrezas más simples, pasibles de cierto grado de automatización mediante la práctica, no implica que este complejo mayor posea estas características.

En el libro *Los contenidos de la Reforma*, de Coll y otros (1992), los autores presentan un cuadro en el que clasifican las estrategias cognitivas para el aprendizaje. Si bien este libro en su totalidad, así como estos intentos clasificatorios en particular, nos inducen a desarrollar ciertas actitudes de prevención y cuidado por el peligro de esquematización que representan, tomaremos de allí algunos ejemplos a fin de facilitar la comprensión de este concepto para quienes no estén familiarizados con él.

- Habilidades en la búsqueda de información (cómo encontrar dónde está almacenada la información respecto a una materia; cómo hacer preguntas; cómo usar una biblioteca, etcétera).
- Habilidades de asimilación y retención de la información (cómo escuchar para la comprensión; cómo estudiar para la comprensión; cómo recordar, cómo codificar y tomar representaciones, etcétera).
- Habilidades organizativas (cómo establecer prioridades; cómo disponer los recursos; cómo conseguir que las cosas más importantes estén hechas a tiempo, etcétera).

- Habilidades inventivas y creativas (cómo desarrollar una actitud inquisitiva; cómo razonar inductivamente; cómo generar ideas, hipótesis, predicciones; cómo organizar nuevas perspectivas; cómo utilizar analogías, etcétera).
- Habilidades analíticas (cómo desarrollar una actitud crítica; cómo razonar deductivamente; cómo evaluar ideas e hipótesis, etcétera).
- Habilidades en la toma de decisiones (cómo identificar alternativas; cómo hacer elecciones racionales, etcétera).
- Habilidades de comunicación (cómo expresar ideas oralmente y por escrito).
- Habilidades sociales (cómo cooperar y obtener cooperación; cómo competir lealmente, etcétera).
- Habilidades metacognitivas (cómo evaluar la ejecución cognitiva propia; cómo seleccionar una estrategia adecuada para un problema determinado; cómo determinar si uno comprende lo que está leyendo o escuchando; cómo transferir los principios o estrategias aprendidos de una situación a otra, etcétera).

La evaluación de las E. de A. se acerca al concepto de metacognición. Novak y Gowin (1988) definen dos conceptos cercanos: *metaconocimiento y metaaprendizaje: "Por metaconocimiento se entiende el conocimiento relativo a la naturaleza del conocimiento y del conocer". "El metaaprendizaje se refiere al aprendizaje relativo a la naturaleza del aprendizaje; es decir, aprendizaje sobre el aprendizaje."*

Desde allí, entendemos que la evaluación de las E. de A. consiste en referir los datos suministrados por la evaluación a los procesos y estrategias cognitivas utilizadas para aprender. Es el grado de conciencia que tiene una persona acerca de sus formas de pensar (aprender) y de la estructura de sus conocimientos. O el que es capaz de inferir acerca del conocimiento llevado a cabo por otro.

Su utilidad es tanto para el docente como para los alumnos, porque significa tomar conciencia a partir del análisis evaluativo de:

- cuáles son las formas en que aprende mejor;
- cuándo y por qué aparecen obstáculos y dificultades;
- cómo recuerda mejor;

- cuáles son los dominios de conocimiento que tiene más desarrollados y cuáles menos;
- cuál es el grado de conocimiento que se posee sobre cada E. de A., así como su uso y aplicación pertinente a situaciones particulares.

En una institución educativa, las E. de A. que desarrollan los alumnos parecen estar en íntima relación con las estrategias de enseñanza que ponen en práctica los profesores, pero también con los contenidos disciplinares de las distintas áreas. Hay algunos avances en las investigaciones que confirmarían lo antes dicho. Esto podría significar un aporte importante al campo de la Didáctica.

Sintetizando: ¿qué entendemos por evaluar las E. de A.? Si evaluar es, en primera instancia, una tarea que apunta a conocer y comprender, la evaluación de las E. de A. será una estupenda forma de apropiación de los procesos por los cuales ese conocimiento fue posible. En segundo término, evaluar será también reflexionar y juzgar acerca de la calidad y eficacia de dichas estrategias y la adecuación de su elección al tema/objeto del aprendizaje en cuestión.

¿Quién evalúa las E. de A.? Ante todo, el propio sujeto que las lleva a cabo; pero también el docente puede llegar a conocerlas y ser capaz de emitir sus apreciaciones sobre la base de sus juicios de valor y su postura respecto del conocimiento.

Por ejemplo, un grupo de alumnos puede recurrir, de modo sistemático, a E. de A. de carácter asociativo y repetitivo, ligadas y favorecedoras de un aprendizaje memorístico (repetición, subrayado, copia, etcétera). Promover procesos favorecedores de instancias metacognitivas, que permitan la toma de conciencia y la reflexión crítica sobre los mismos, sus alcances y limitaciones, puede ser el primer paso para su desestructuración y transformación en otras que requieran procesos de organización y elaboración del conocimiento.

En el otro extremo, saber que estamos construyendo y poniendo en práctica estrategias que favorecen el aprendizaje constructivo, con sustento en significaciones personales, basadas en relaciones conceptuales pertinentes, seguramente será un aporte valioso para incentivar y favorecer dicho proceso.

¿Cómo evaluar las E. de A.? Si se trata del propio sujeto, la

evaluación de sus E. de A. consistirá, en primer lugar, en un acto analítico-introspectivo, que le permita recuperar y hacer conscientes los distintos tipos de procesos que pone en juego al aprender y, en segundo lugar, reflexionar sobre su pertinencia y calidad para la apropiación y manejo de esos contenidos, así como su eficacia respecto del tiempo y esfuerzo empleados.

Si se trata de evaluar las E. de A. que otro pone en juego, parecería necesario pensar en una metodología cercana a la observación participante, propia de la etnografía y la antropología social. Es decir, observar, preguntar, registrar, aportar elementos para promover la reflexiones del sujeto.

Recordemos que no se trata de "poner una nota" de acuerdo con las E. de A. que desarrollen los alumnos. Al pretender evaluarlas, estamos pensando en hacer posible su explicitación y, a través de ella, transparentarlas, intentar comprender las razones de su construcción y/o elección como modo de conocimiento.

> h) El uso de la información proveniente de las acciones evaluativas pone de manifiesto el tema del poder en este campo, permitiendo o dificultando, según los casos, la apropiación democrática del conocimiento que en él se produce.

Parece casi obvio que todo acto de evaluación educativa lleva aparejada instancias de información. De hecho, casi no se concibe su realización si no va acompañada de alguna forma de comunicación.

Sin embargo, sólo en apariencia, esto es algo incuestionable y carente de problemas. El tema es qué se informa y para qué. Como veremos en seguida, son varios los aspectos para analizar, y no pocos los momentos en que habrá que tomar decisiones.

Antes de proseguir, desearía explicitar una observación: creo que la relación evaluación-poder es uno de los temas centrales del campo evaluativo de los '90. Autores como House (1994), Popkewitz (1994), Santos Guerra (1993), Angulo Rasco (1994), Cano (1995), Celman (1996, 1997), entre otros, han realizado un desarrollo teórico de las determinaciones mutuas entre ambos, así como han

mostrado, empíricamente, su concreción en diferentes ámbitos y en distintos países.

Aquí intentaremos analizar brevemente esta cuestión, refiriéndola circunscriptamente al ámbito de la evaluación de los conocimientos.

Consideremos para ello dos casos concretos y opuestos. Uno, el de aquel profesor que dispone por sí, no sólo el momento, la forma y el contenido de las pruebas específicas de evaluación de los aprendizajes de sus alumnos, sino que, además, las corrige sin explicitar los criterios por los cuales ha juzgado correcto o incorrecto, adecuado o inadecuado, el trabajo realizado, comunicándoles, finalmente, sólo los resultados obtenidos.

¿Qué aprenden y qué dejan de aprender los estudiantes en una situación así?

- Aprenden que el conocimiento es un proceso que no les pertenece, que se realiza "para otros", quienes tienen la posibilidad de decidir; y dejan de aprender que existen diversos criterios, no sólo los que ha utilizado el docente; que lo importante es explicitarlos y fundamentarlos, porque esto también es aprendizaje.
- Aprenden que no tienen derecho a conocer los fundamentos y razones del juicio que ha emitido su profesor sobre su propia tarea, porque ellos "son sólo alumnos", "no saben de eso"; no están capacitados; y dejan de aprender, de ejercitarse, en el uso de sus derechos como sujetos, como ciudadanos y en el desarrollo de un razonamiento que les permita emitir juicios justificados...
- Aprenden cuáles son las cosas que con mayor probabilidad los pueden ayudar a obtener buenas notas o, cuanto menos, aprobar con ese docente, realizando aquello que han descubierto que les agrada; y dejan de aprender qué valor tiene lo que aprenden, qué es lo que no saben, cómo es que no lo saben, qué otro conocimiento los puede ayudar, qué deberían hacer para saber, etcétera.
- Aprenden a ser dependientes y poner fuera de ellos la responsabilidad sobre lo que les ocurre con sus propios procesos de estudio y conocimiento, esperando la aprobación o desaprobación

externa para, recién entonces, poder emitir su opinión al respecto, y dejan de aprender las herramientas necesarias para poder decidir, fundadamente, en cuestiones que van más allá de lo escolar.

Libertad y autonomía, democracia y autovalía son también cosas que se aprenden en las escuelas. Y el espacio de las prácticas evaluativas es, entre otros, un lugar potente para esos aprendizajes. O los contrarios.

El otro caso es el del docente que acuerda con el grupo las razones y finalidades de una determinada actividad de desarrollo y explicitación del aprendizaje, el momento, la forma y los contenidos que abarcará, las formas y criterios con que podrá analizarse y las derivaciones que pueden efectuarse a partir de su evaluación.

Barry MacDonald (1985), desarrollando el área de la evaluación institucional y de proyectos, sostiene que el trabajo del evaluador es eminentemente político y sus diversos estilos y métodos son la expresión de diferentes actitudes en relación a la distribución del poder en educación. Eligiendo sus fidelidades y prioridades, el evaluador necesariamente adopta una postura política.

Acordando con esta postura, creemos que el profesor que adopta una concepción democrática de la evaluación prioriza a los alumnos –sin descartar absolutamente a otros– como los primeros sujetos con derecho a participar en los procesos relacionados con su aprendizaje y, por ende, también a conocer la información en torno de él.

Esto implica, dentro del aula, que deberá cuidar, entre otras cosas, las acciones que continúan luego de una instancia evaluativa. Dedicar tiempo a las instancias de "devolución de la información", de modo que ese diálogo facilite la comprensión de los factores intervinientes en el transcurso de la enseñanza y el aprendizaje. Asimismo, intercambiar ideas acerca de las posibilidades de mejora es un aspecto fundamental en el intento de convertir a la evaluación en herramienta del conocimiento.

Una última advertencia al respecto: los docentes que se decidan a transitar este camino deben saber que están abriendo espacios para la crítica, el cuestionamiento y el juicio, y también respecto de su propia tarea. Esto implica, entonces, abandonar un lugar seguro y tranquilo, que es el del evaluador con la suma del poder. Pero como

contrapartida, es ganar colaboradores en este complejo trabajo de intentar formular las preguntas que nos permitan comenzar a comprender "qué está sucediendo aquí" (Santos Guerra, 1990).

Conclusiones a modo de nuevas aperturas

En este capítulo hemos tratado de presentar, implícitamente, un enfoque de la evaluación educativa alejado de la constatación, la medición y la comparación (¿competitiva?) de los conocimientos.

B. Bernstein dice que el ritmo de los aprendizajes que se exigen a los escolares es hoy tan intenso que no basta el tiempo de la escuela para tener éxito en su empeño. Por eso hace falta una segunda escuela (la casa) para conseguirlo. Santos Guerra (1990) se pregunta al respecto de estas palabras de Bernstein: *"¿Qué sucede con los que no tienen ese segundo lugar facilitador de la tarea? ¿No es cierto que vuelven a ser machacados por las exigencias del sistema (en este caso llamado paradójicamente educativo)?"*.

No nos parece demasiado justo el uso de procedimientos iguales entre desiguales, aunque aparentemente lo sea y algunos estén interesados en que así se los considere.

En confrontación con esta posición es que nos preguntamos si era posible mejorar la evaluación y transformarla en herramienta de conocimiento. Como puede observarse a esta altura del desarrollo del tema, nuestra respuesta es afirmativa. Pero, como mínimo, deberán cumplirse dos condiciones:

- *condición de intencionalidad*. Para utilizar la evaluación como un modo de construcción de conocimiento fundado, autónomo y crítico, los sujetos deben estar interesados en ello. Estas cuestiones raramente despiertan adhesiones irracionales. Son, generalmente, el producto de un trabajo reflexivo y consciente, asentado en posturas éticas. Pero, como unas y otras no son innatas sino que se construyen personal y socialmente, dejamos estas puertas abiertas a la consideración de quienes estén dispuestos a repensarlas y recrearlas; y
- *condición de posibilidad*. Por tratarse de una propuesta que no consiste en seguir un plan previsto, detallado y preciso, sino en

el ofrecimiento de unas herramientas para un trabajo artesanal cuyo modelo, si existe, lo crea el propio sujeto, la evaluación entendida en estos términos requiere:

- que los sujetos se sientan tales, es decir, que puedan desplegar cierto grado de autonomía, autoestima y autovalía personal;
- que exista, o se genere, un medio educativo que admita o, mejor aún, valore estas actividades, y
- que se creen las condiciones institucionales y materiales de trabajo docente para su desarrollo.

Una escuela que no esté dispuesta a exponerse al juego democrático, probablemente no adoptará estos principios.

Pero unos docentes que estén sometidos a condiciones de trabajo tales que no cubran los requisitos básicos y mínimos para el desarrollo de su tarea, quizá deban hacer uso de ellos, en primera instancia, para avanzar en el conocimiento de su propia situación de educadores y utilizarlos también como herramientas para promover críticamente su profesionalidad y la autonomía de sus alumnos.

Bibliografía citada

Allal, L.: "Estrategias de evaluación formativa. Concepciones psicopedagógicas y modalidades de aplicación", *Infancia y aprendizaje*, nº 11, Madrid, 1980.

Álvarez Méndez, N. M.: *Valor social y académico de la evaluación*, mimeografiado, Madrid, 1993.

———: "El alumnado. La evaluación como actividad crítica del aprendizaje", *Cuadernos de Pedagogía*, nº 219, Barcelona, 1993.

Angulo Rasco, F.: *Innovación y Evaluación educativas*, Málaga, Universidad de Málaga, 1990.

———: "¿A qué llamamos evaluación?: Las distintas acepciones del término 'evaluación' o por qué no todos los conceptos significan lo mismo", en J. F. Angulo y N. Blanco (comp.): *Teoría y desarrollo del currículum*, Málaga, Ediciones Aljibe, 1994.

Ausubel, D. P.; Novak, J. D., y Hanessian, M.: *Psicología Educativa. Un punto de vista cognoscitivo*, México, Trillas, 2ª edición, 1986.

Bertoni, A. y otros: *Evaluación. Nuevos significados para una práctica compleja*, Buenos Aires, Kapelusz, 1996.

Cano, D.: *Evaluación institucional en las universidades argentinas: una aproximación bibliográfica a una historia reciente*, Mejoramiento de la calidad académica de la UNL, Universidad del Litoral, 1994.

———: "Principios y criterios rectores para una política de evaluación institucional", *La evaluación institucional*, Secretaría Académica, UNL, 1995.

Celman, S.: *Evaluación de proyectos institucionales. Dime por qué preguntas y te diré quién eres*, Congreso Internacional de Formación de Profesores, UNL, Santa Fe, 1996 (en prensa).

———: *La evaluación universitaria en un contexto de ajuste*, Reunión Científica sobre Evaluación Institucional de las Universidades. Eficacia o calidad, Almería, Universidad de Almería (en prensa).

Coll, C. y otros: *Los contenidos de la Reforma. Enseñanza y aprendizaje de conceptos, procedimientos y actitudes*, Madrid, Santillana, 1992.

House, E. R.: *Evaluación, ética y poder*, Madrid, Morata, 1994.

MacDonald, B.: "La evaluación y el control de la educación", en Gimeno Sacristán, J. y Pérez Gómez, A.: *La enseñanza: su teoría y su práctica*, Madrid, Akal/Universitaria, 1985.

Monereo, C. y otros: *Estrategias de enseñanza y de aprendizaje. Formación del profesorado y aplicación en la escuela*, Barcelona, Grao, 1994.

Newman, D.; Griffin, P., y Cole, M.: *La zona de construcción del conocimiento*, Madrid, Morata, 1991.

Nisbet, J. y Shuckmith, J.: *Estrategias de aprendizaje*, Madrid, Santillana, 1987.

Novak, J. y Gowin, D.: *Aprendiendo a aprender*, Barcelona, Martínez Roca, 1988.

Pérez Gómez, A.: "La sociedad posmoderna y la función educativa de la escuela", en *Escuela pública y sociedad neoliberal*, Málaga, Edit. Aula libre, 1997.

Popkewitz, T.: *Sociología política de las reformas educativas*, Madrid, Morata, 1994.

Pozo, J. y Postigo Aragón, Y.: *Las estrategias de aprendizaje como contenidos del curriculum*, mimeografiado, Barcelona, 1993.

Santos Guerra, M. A.: *Hacer visible lo cotidiano. Teoría y práctica de la*

evaluación cualitativa de los centros escolares, Madrid, Akal/Universitaria, 1990.

— *La evaluación: un proceso de diálogo, comprensión y mejora,* Málaga, El Aljibe, 1993.

riven, Michael: "The methodology of evaluation", en Stake, R.: *Perspectives of Curriculum Evaluation,* American Educational Research Association, Monograph Series on Curriculum Evaluation nº 1, Chicago, Rand McNally, págs. 38-39, 1967.

Stufflebeam, D. y Shinkfield, A.: *Evaluación sistemática. Guía teórica y práctica,* Barcelona, Paidós, 1987.

LA CALIDAD DE LOS PROGRAMAS DE EVALUACIÓN Y DE LOS INSTRUMENTOS QUE LOS INTEGRAN*

Alicia R. W. de Camilloni

A Tomás y Carolina

INTRODUCCIÓN

La evaluación de los aprendizajes se realiza sobre la base de un programa que, al servicio de la enseñanza y del aprendizaje, está constituido por un conjunto de instrumentos de evaluación. Dado que cada tipo de instrumento permite evaluar diferentes aspectos de los aprendizajes de los alumnos, es menester garantizar la pertinencia y calidad técnica del programa considerado integralmente como una estructura, así como la de cada uno de sus componentes.

Un principio general que rige el diseño de un programa de evaluación es el de la economía de tiempo. El tiempo destinado a situaciones de evaluación puntual, diferenciadas de las situaciones de enseñanza, debe ser el más breve posible, ya que el factor más crítico de la educación formal es la carencia de tiempo suficiente para enseñar con profundidad toda la información y para desarrollar todas las competencias que los alumnos deben alcanzar.

Las condiciones en que se desenvuelve actualmente la educación escolar en todos los niveles implican, precisamente, la necesidad de sortear algunas dificultades para la evaluación porque las clases suelen ser numerosas como producto de los muy significativos y deseados procesos de democratización de la educación. El número de alumnos ha aumentado en los niveles de educación básica,

* Este artículo ha sido posible gracias al subsidio otorgado a la investigación dirigida por la autora por la Secretaría de Ciencia y Técnica de la UBA (FI215).

secundaria y superior, planteando el requerimiento de encontrar las técnicas adecuadas para la enseñanza de grupos numerosos, así como las técnicas de evaluación que puedan acompañar adecuadamente a una enseñanza que promueva aprendizajes significativos en todos los alumnos, que evite el fracaso escolar, la repitencia y la deserción, y que permita hacer efectivo el propósito de igualar las oportunidades educativas reconociendo los derechos a la igualdad así como a la diversidad de los rasgos personales de los alumnos.

Por ello, la calidad de un programa de evaluación está sujeta, desde un punto de vista pedagógico y por esto, también, ético y político, a su capacidad para evaluar justa y equitativamente a grupos numerosos integrados por alumnos diferentes.

No es posible hablar de la evaluación de los aprendizajes al margen de los procesos de enseñanza y aprendizaje que los han generado. Una condición fundamental para el diseño de programas de evaluación, en cualquiera de los niveles que se deseen o se deban evaluar en la institución escolar, es que tienen que ser consistentes con los proyectos de enseñanza y aprendizaje de la institución ya que sólo alcanzan su propósito de servir como dispositivos para el perfeccionamiento de los resultados de la educación cuando se convierten en juicios de autoevaluación, tanto para los alumnos como para los docentes y las autoridades de la escuela y del sistema.

Para lograr una buena configuración de los procesos de enseñanza, el docente debe tomar algunas decisiones adecuadas. Las decisiones de los docentes están fundadas sobre las concepciones que ellos tienen, entre otras, acerca de qué es enseñar, qué es aprender, cuál es la naturaleza de los conocimientos que los alumnos deben adquirir y qué funciones tiene que cumplir la escuela. Hay, pues, un conjunto de teorías o ideas más o menos explícitas, o a veces, implícitas, basadas en el sentido común o en teorías científicas, según los casos, a través de las cuales cada docente procura construir explicaciones de los procesos y las funciones relacionadas con la intervención pedagógica. Esas explicaciones le sirven de fundamento cuando diseña su proyecto de enseñanza.

Los docentes también ponen en juego la información que poseen acerca de lo que está ocurriendo con sus alumnos y las características que suponen que éstos tienen como individuos y como grupo de clase, ya que ellos son los destinatarios de la enseñanza. En esta

tarea, la referencia se centra en la necesidad de prever cuáles habrán de ser los impactos más duraderos y los efectos más inmediatos que posiblemente tengan sus propias acciones de enseñanza sobre los alumnos. Para efectuar esas previsiones, el docente debe poseer información de distintos tipos. Ella puede ser recogida por medio de la observación del desarrollo de los procesos de enseñanza o puede surgir de las respuestas de los alumnos colocados en situaciones creadas especialmente por los docentes con ese fin.

Si se parte de estas premisas, se advierte que la importancia del papel que desempeñan las teorías explicativas de cómo ocurren estos procesos, qué modalidades adoptan en su presentación y cuáles son las variables que inciden en ellas, obliga al docente a efectuar el análisis y la reconstrucción de dichas explicaciones teóricas, de modo de trabajar con aquellas a las que estima más apropiado adjudicar un mayor valor científico por su actualidad, seriedad y rigor.

Pero, además, en lo que se refiere a la información relativa a los individuos concretos o empíricos que conforman el grupo de clase o de la institución, el docente necesita contar, también, con la mejor información. Para obtenerla puede recogerla de las situaciones naturales tal como ocurren en la sala de clase o fuera de ella, o bien puede generar intencionalmente situaciones específicas. La evaluación tiene un estado inicial que puede construirse a partir de dos tipos de operación: una modalidad se constituye sobre la base de recoger datos a través de la observación sistemática o asistemática o de algún otro medio no diferenciado del proceso de enseñanza volcados en la elaboración de sus respectivos registros y, la otra, con características bien diferenciadas de la anterior, consiste en obtener datos en situaciones especialmente diseñadas para recoger una información que se considera relevante como indicadora del estado de situación de los aprendizajes que se quieren evaluar.

Además de esta operación inicial, en alguno de sus dos tipos básicos, para que haya una evaluación con sentido es indispensable que existan criterios que permitan construir juicios de valor acerca de lo que la información recogida significa en términos de aprendizaje de los alumnos. La primera operación, por lo tanto, consiste en recoger información; y la segunda, en analizar e interpretar la información recogida sin que, necesariamente, el orden enunciado

implique la diferenciación temporal o el establecimiento de una secuencia de estas dos operaciones. La interpretación permite completar el proceso de evaluación y conduce a la construcción de juicios de valor que son característicos de la evaluación como función didáctica. No hay evaluación sin juicios de valor. Debemos mencionar aquí, sin embargo, que algunos autores discrepan con esta afirmación. Jean Cardinet (1989), por ejemplo, advierte, como veremos más adelante, acerca de los riesgos que implica el uso de juicios de valor en la evaluación.

El juicio de valor puede ser construido de dos maneras. Según una de ellas, puede responder a estados afectivos del docente, relacionados con actitudes de aceptación o rechazo, de agrado o desagrado ante determinada conducta del alumno, observada directamente, inferida de conductas observadas o, también, sólo supuesta más o menos indirectamente a partir de la interpretación de variadas situaciones. El juicio de valor puede surgir, igualmente, como producto de una elaboración seria y rigurosa de la información recogida sistemáticamente, a partir ya no de emociones, sentimientos o actitudes, sino de una base de conocimientos que permite fundamentar el juicio de valor de una manera que pretende ser "objetiva". Ante estas dos maneras diferentes de construcción es posible que, por la naturaleza misma de los actos pedagógicos, en el ejercicio integral del papel del docente se requiera poner en juego las dos modalidades de construcción de los juicios de valor, buscando el modo de articularlas adecuadamente, entendiendo que el docente es una persona y que la relación educativa es una relación entre personas.

La evaluación tiene que contar con criterios que permitan analizar e interpretar la información recogida con base en una teoría orientada a responder a la serie de cuestiones que antes mencionamos. Entre ellas incluimos la teoría referida a la evaluación como función didáctica. Si el docente tiene que contar con teorías explicativas acerca de cómo aprenden los alumnos y sobre qué es enseñar, también deberá trabajar con teorías acerca de la evaluación, su relación con los procesos de enseñanza y aprendizaje, el lugar que ocupa la evaluación en la administración de un proyecto de enseñanza, qué alcances y qué límites debe tener.

Es siempre importante insistir en la necesidad de que todas esas teorías sean consistentes entre sí y conformen, en su conjunto, una

base de percepción, pensamiento, sentimiento y conocimiento que permita fundar y justificar la acción docente.

Estos principios no suponen, sin embargo, que exista una única concepción científica acerca de las cuestiones que hemos mencionado, incluyendo entre ellas a la propia evaluación. Así como hay diversas concepciones acerca del universo, las hay también sobre la enseñanza y el aprendizaje.

Cuando se adopta una concepción acerca de la enseñanza, de lo que la enseñanza debe ser y no sólo acerca de lo que ella es, esta concepción normativa conduce a una idea normativa de la evaluación, de lo que la evaluación debe ser. Así pues, para una teoría normativa de la enseñanza, no todas las concepciones sobre la evaluación son válidas; algunas son aceptables, otras no lo son.

La evaluación es un recurso indispensable para el perfeccionamiento de los dos procesos que nos interesan, la enseñanza y el aprendizaje. Encarar una acción docente significa actuar para mejorar las probabilidades de que se produzcan procesos deseables de aprendizaje. Pero es menester señalar que, aunque por lo general cuando hay enseñanza se produce algún aprendizaje, éste no siempre coincide con los propósitos del docente. La actitud del docente frente a la evaluación deberá ser abierta, coincidiendo con su teoría de la enseñanza. Si se está de acuerdo con la idea de que, al enseñar, el docente no debe desarrollar una intervención caracterizada por su unidireccionalidad en la que la única voz a escuchar es la del propio docente sino que hay que dar lugar a la voz del alumno, esto es, a la manifestación de su capacidad para pensar y construir significados, del mismo modo en el proceso de evaluación debe encontrar el alumno un lugar para expresar los significados desde su propia perspectiva.

Cuando se opta, entonces, por concepciones de la enseñanza en las que se manifiesta nuestro respeto por la capacidad de producción personal del alumno, los instrumentos de evaluación que utilicemos deberán ser coherentes con esa postura.

En la actualidad disponemos de una cantidad importante de instrumentos diversos, creados muchos de ellos a lo largo del siglo XX. Consisten, por un lado, en las llamadas "pruebas objetivas", diseñadas a partir del programa de investigación científica del conductismo, con una serie de especificaciones técnicas importan-

tes en cuanto a la construcción de ítemes y al análisis e interpretación de resultados. También contamos con un conjunto de normas técnicas para la utilización de las denominadas "pruebas subjetivas"–a veces llamadas "tradicionales"– en las que, a diferencia de las pruebas objetivas que sólo admiten una única respuesta correcta, se aceptan diferentes niveles de calidad de respuesta. En este último campo, las normas técnicas desarrolladas han apuntado a mejorar la construcción, la administración y la interpretación de los resultados, con el propósito de mejorar también la confiabilidad y la validez de las pruebas subjetivas.

En este siglo, el trabajo realizado en el campo de la evaluación didáctica ha sido muy profuso, ya que se ha considerado que se trataba de un tema prioritario en las teorías de la enseñanza. No es extraño, por ejemplo, que para las diversas teorías conductistas del aprendizaje, la evaluación se convirtiera en un aspecto central de su desarrollo teórico y experimental y de su propuesta didáctica, hasta el punto de sostener que se debía limitar lo que se enseñaba a aquello que era posible evaluar en los términos que la teoría aceptaba con carácter exclusivo, es decir, cuantificando los rasgos de las conductas observables. Dada la significativa importancia que se otorgaba a la evaluación de los aprendizajes y a la simplicidad técnica del problema planteado, se comprende la gran fertilidad de las teorías conductistas en el dominio teórico y práctico de la evaluación.

Consideramos que se trata de una función fácil de resolver técnicamente en el marco de las teorías del conductismo porque se limita a evaluar lo que se puede observar. Las corrientes cognitivas, en cambio, se proponen evaluar los procesos mentales. Su problema es más difícil de resolver que el que se planteaban los conductistas y sus respuestas técnicas son más dificultosas de construir. La cuestión de la evaluación, en cambio, es menos importante para las didácticas humanistas que plantean una relación diferente entre docente y alumno en la que los aprendizajes son aún más difíciles de evaluar que en las propuestas didácticas cognitivas.

Debemos señalar, empero, que el programa conductista caracterizado por su gran capacidad de producción de instrumentos técnicos de evaluación durante varias décadas del siglo XX no fue el único programa que realizó trabajos en este tema; no fue ni siquiera el que ejerció una hegemonía completa en la evaluación. En el

campo teórico y práctico, simultáneamente con el trabajo que en este ámbito fuera realizado por los psicólogos conductistas, también se trabajó con otros instrumentos de evaluación que respondían a perspectivas diferentes. Algunas corrientes europeas que desarrollaron trabajos importantes en este terreno lo hicieron desde enfoques diferentes, alternativos a las producciones conductistas. Entre ellos, por ejemplo, se destacan los que entregan al alumno la responsabilidad de seleccionar el material, organizarlo y encontrar formas adecuadas de expresar su producción personal, cuestiones que no pueden ser evaluadas por medio de la administración de pruebas objetivas. De este modo, y aún en el ámbito de mayor influencia de las teorías conductistas, se emplearon siempre y se enseñó a los docentes a construir, analizar e interpretar los resultados de las llamadas "pruebas subjetivas" o "tradicionales".

Es importante hacer notar que la gran disponibilidad de instrumentos de diferente carácter, alcance y función, es un factor de enriquecimiento del abanico de posibilidades que se abren para el diseño de programas de evaluación. El problema que se ha de resolver es el de su adecuada selección y organización de modo de alcanzar consistencia con el proyecto docente.

Uno de los productos más importantes del trabajo de los diversos programas de investigación sobre la evaluación en este siglo ha sido, precisamente, el mejoramiento de la normativa de construcción de instrumentos y análisis de resultados, lo cual nos permite abordar el problema de la evaluación contando con una amplia variedad de instrumentos que incluyen pruebas escritas y orales y pruebas de observación de procesos y de producciones, utilizables de diversas maneras según las necesidades de la enseñanza. El aporte del conocimiento desarrollado en las últimas décadas sobre los procesos cognitivos da un nuevo sustento a las teorías y técnicas evaluativas y permite confirmar o modificar las prácticas instaladas en la educación formal desde hace largo tiempo en algunos casos, y en otros recientemente, como resultado de la introducción de nuevas propuestas que es necesario reanalizar a la luz de la investigación actual.

Los programas de evaluación

Si nos preocupamos hoy por determinar cuáles son los ejes en torno de los cuales gira esta cuestión, tendríamos que comprender, en primer lugar, que el propósito principal de la evaluación no se puede lograr si la evaluación no se convierte en autoevaluación tanto para el docente cuanto para el alumno. La evaluación, en segundo lugar, debe ser consistente con las concepciones de la enseñanza y del aprendizaje. En tercer lugar, la resolución técnica que se le dé debe permitir evaluar todos los aspectos que están comprometidos con los procesos de aprendizaje, lo cual supone desarrollar programas complejos de evaluación en los que se empleen una cantidad de instrumentos diversos y donde cada técnica sea adecuada para evaluar diferentes aspectos.

Los programas de evaluación deben ser diseñados por los docentes, como parte de la programación didáctica de sus cursos. Para ello, el docente debe conocer en profundidad la o las teorías de la evaluación y la variedad de instrumentos que existen. Estos conocimientos le permitirán diseñar programas coherentes con la programación de la enseñanza combinando instrumentos diversos, así como crear nuevos instrumentos de evaluación acordes con la multiplicidad y heterogeneidad de los propósitos de su enseñanza. Es también indispensable, si quiere desarrollar un buen programa de evaluación, que el docente conozca las normas técnicas para la construcción, administración, análisis e interpretación de resultados.

Queremos señalar que, en la medida en que un docente es responsable de la elaboración de su programación didáctica, deberá tener la misma libertad para diseñar y administrar su programa de evaluación. Los grados de libertad en la toma de decisiones deben ser equivalentes. Si así no ocurriera y si el docente tuviera libertad para decidir algunos aspectos de su proyecto de enseñanza y no la tuviera para decidir acerca de su programa de evaluación porque ésta le es impuesta desde el exterior, la evaluación no cumpliría otra función que la de servir de control, resultando minimizados sus posibles efectos positivos de mejoramiento de los procesos y de los resultados de esos procesos de enseñanza y de aprendizaje. Hay que recordar siempre que la libertad de quien enseña debe tener su correspondencia en la libertad de quien evalúa y en la libertad de

quien aprende. Y, respecto de este último es importante recordar, como lo hace E. R. House (1994) citando a John Rawls, que "... respetar al otro como persona moral supone tratar de comprender sus aspiraciones e intereses desde su punto de vista y presentarle consideraciones que le hagan posible aceptar las limitaciones de su conducta" (pág. 129).

Los determinantes de la calidad de la evaluación

Un programa de evaluación y cada uno de los instrumentos de evaluación que lo integran deben cumplir con ciertas condiciones. Mencionaremos aquí algunas de ellas: debe permitir evaluar los aprendizajes que la institución y el o los docentes esperan que los alumnos desarrollen. Las consignas de trabajo que integran el instrumento deben ser seleccionadas, en consecuencia, de modo de permitir que tanto la presencia de una respuesta como su ausencia sean significativas en relación con los aprendizajes que se ha buscado promover. En términos generales, las consignas no deben producir respuestas que resulten irrelevantes para la evaluación de los aprendizajes previstos. Esto no significa, sin embargo, que se descarten las respuestas irrelevantes de los alumnos cuando aparecen, ya que a través de esos resultados en apariencia disfuncionales, a veces incomprensibles, el docente puede advertir errores que sí pueden ser altamente significativos y que requieren especial atención. El instrumento, aunque siempre haga una lectura parcial de lo aprendido por los alumnos, debe proponerse registrar no sólo los éxitos y fracasos sino también el origen de esos fracasos. El instrumento debe presentar el grado de organización suficiente para que la apreciación que efectúa del aprendizaje permita desprender algunas conclusiones acerca del desempeño presente y futuro del alumno, en cuestiones específicas pero también con visión integral.

Cada uno de los instrumentos de evaluación resuelve estos problemas de diversas maneras. Actúa como un reflector que en el escenario ilumina a algunos personajes y deja en la penumbra a otros, que sin embargo están allí, presentes, pero con una clase diferente de presencia. Por esta razón, la elección de los instrumentos de evaluación adecuados a la hora de diseñar el programa de

evaluación de un curso, por ejemplo, constituye una de las decisiones más importantes para garantizar el valor didáctico de la evaluación en relación con los procesos de enseñanza y de aprendizaje.

Existe una variedad de modelos de instrumentos de evaluación. Cada uno de ellos, tomado aisladamente, es insuficiente para obtener información sobre el aprendizaje de los alumnos. El diseño del programa exige, pues, la combinación de instrumentos diversos para obtener una cobertura adecuada. La eficacia de la evaluación depende, entonces, de la pertinencia de la combinación de diferentes instrumentos, de la oportunidad en que se administran y de la inteligencia y propiedad del análisis e interpretación de sus resultados.

Veamos, entonces, los criterios o requisitos que deben tenerse en cuenta en el momento de tomar decisiones acerca del programa de evaluación de un curso, por ejemplo, y sobre cada uno de los instrumentos de evaluación que lo integran. Las características generales que los instrumentos de evaluación deben reunir, aunque con distinto grado de relevancia, son cuatro: validez, confiabilidad, practicidad y utilidad.

Validez

Se dice que un instrumento de evaluación es válido cuando evalúa lo que se pretende evaluar con él. Como un instrumento es utilizado para apreciar ciertos logros de aprendizaje de un cierto grupo de alumnos en una cierta circunstancia y en un cierto momento de su proceso de aprendizaje (al inicio, en el transcurso o al final), la validez de un instrumento no puede ser determinada de manera absoluta, sino siempre en relación con su adecuación a los propósitos y situación específica de su aplicación. Como lo señala Henry E. Garrett (1966), "un test es válido para un propósito particular o en una situación especial; no es universalmente válido". Cuando se quiere determinar si un instrumento es válido se requiere, entonces, información acerca de los criterios que han presidido su construcción y su administración. Los criterios son, entonces, externos a la evaluación misma. Esta cuestión sería relativamente simple de resolver si todos los propósitos de la enseñanza, en todos

los niveles de decisión y para todos los actores, fueran coincidentes. Pero rara vez ocurre esto. A veces los objetivos del docente coinciden plenamente con los de la escuela y los de ésta con los de la jurisdicción y los de esta última con los de la nación. Pero la plena armonía de todos esos niveles parece más improbable que posible de alcanzar y, más aún, ni siquiera se puede afirmar que, en lo pormenorizado de los aprendizajes escolares, sea absoluta y efectivamente deseable. Las diferencias en la educación surgen de la acción y la resolución de cada docente y de cada alumno. Se podrán encontrar, en consecuencia, criterios generales, pero también criterios diferentes. Y no tiene sentido hablar de la validez de un instrumento si no se ha validado el criterio. Para lo cual se requiere, a su vez, una validación mediante criterios más amplios. Según Robert L. Ebel (1963), "esto requiere un criterio para validar el otro criterio y así sucesivamente *ad infinitum*".

En la Didáctica, ese proceso de validación al infinito podría ser reemplazado por una teoría de bases firmes, tanto desde un punto de vista científico como filosófico. De otra manera, y ante la obligación ineludible de evaluar a sus alumnos, los docentes se encontrarían en la situación de ser incorporados como nuevos personajes, junto a Aquiles, la tortuga y la flecha, de las aporías de Zenón de Elea.

No se trata, en consecuencia, de validar instrumentos y criterios como en un juego de permanente retroceso hacia los principios, sino de apoyarse en principios didácticos fuertes, con cierto grado de generalidad y que permitan fundamentar racionalmente las decisiones de diseño del programa y de los instrumentos. Veámoslo en algunos ejemplos.

Si para un docente el propósito principal de su enseñanza ha sido procurar que sus alumnos aprendan a resolver un cierto tipo de problemas, el instrumento de evaluación válido será aquel que introduzca problemas al alumno para que éste los resuelva. Deberá tener presente, sin embargo, que no obtendrá la misma información acerca de lo que sus alumnos saben cuando en el instrumento se plantean problemas que corresponden sólo al tipo de problema cuya solución los alumnos deben hallar, que si también incluyen otros problemas de tipos diferentes, de modo que los estudiantes deban reconocer a qué tipo de problema corresponde cada uno de

los planteados por el docente. Se podría afirmar que, en tanto el propósito didáctico no habría de ser que el alumno resuelva sólo una clase de problemas, la capacidad de discriminar clases de problemas debería estar presente en la formulación de los propósitos de un programa de enseñanza, por lo que esa segunda alternativa, que incluye diversos tipos de problemas en el instrumento, es más válida que la primera, que presenta solamente un único tipo de problema. Así, pues, para lograr que la construcción del programa o de un instrumento de evaluación sean válidos, es menester poner en juego no sólo una lectura fragmentada de los propósitos de la enseñanza sino principios didácticos más generales.

En otro sentido y refiriéndonos a otro ejemplo, veamos el papel que tiene el tipo de instrumento: si se trata de construir un examen oral o un coloquio válidos, hay que definir previamente preguntas centrales. Éstas deben referirse a cuestiones fundamentales del dominio de conocimiento que el alumno debe aprender. Pero la validez no llega a estar asegurada de este modo ya que en estos tipos de instrumentos de evaluación, el examen oral y el coloquio, la característica que les es propia y que hay que desarrollar adecuadamente es la elaboración de preguntas o comentarios del docente a partir de las respuestas del alumno. Si ellas no le permiten al alumno corregir sus respuestas anteriores, por ejemplo, o no lo animan a profundizar respuestas que inicialmente eran superficiales, o si no lo inducen a la elaboración de asociaciones y conclusiones reuniendo respuestas fragmentarias, la técnica del examen oral o del coloquio están mal empleadas y podrían ser reemplazadas por un instrumento de respuesta escrita. Como se ve, la validez de los instrumentos también depende de su apropiada construcción en relación con las características específicas que los definen.

Se puede observar, a través de lo expuesto hasta aquí, que la validez nunca es absoluta. En primer lugar, porque siempre debe ser referida a un criterio externo. En segundo lugar, porque sólo se puede hablar de validez mayor o menor, ya que un instrumento nunca es "completamente" válido pues sólo permite hacer inferencias acerca de los saberes de los alumnos. Y finalmente, porque la validez está condicionada a un ajuste que nunca puede ser perfecto entre el saber aprendido que se quiere evaluar y los medios técnicos, esto es, los instrumentos disponibles para evaluarlo. Con

frecuencia se obtiene mayor validez a través de la combinación feliz de distintos tipos de instrumentos.

Se diferencian diversas clases de validez. Desarrollaremos sintéticamente sus principales modalidades.

Validez de contenido

Se dice que un instrumento tiene validez de contenido cuando representa una muestra significativa del universo de contenido cubierto por el curso o la unidad didáctica, la clase o una serie de clases. La validez de contenido, por lo tanto, requiere que la selección de contenidos que se ha efectuado para construir el instrumento sea equivalente a la de cualquier otro conjunto de contenidos que se hubiera podido seleccionar con ese fin. Se trata, por supuesto, en todos los casos de contenidos importantes por su significación en el aprendizaje de los alumnos.

La dificultad que la construcción de un instrumento debe superar es la necesaria distancia que existe entre la gran cantidad y variedad de los contenidos de toda enseñanza y la brevedad exigida también para los tiempos destinados a una evaluación de los aprendizajes, si no se quiere fatigar a los alumnos, lo cual conduciría a desvirtuar el significado cognitivo de sus respuestas.

Esa distancia obliga al docente a elegir los temas, informaciones, competencias, habilidades, destrezas, procedimientos u otros que incluirá en el instrumento de evaluación. Cuanto mayor es la diferencia entre tiempos destinados a la enseñanza y al estudio y tiempos destinados a la evaluación, más difícil es lograr una validez de contenido aceptable, es decir, que permita asegurar que los resultados obtenidos por el alumno en una instancia de evaluación serían semejantes a los que obtendría en otra distinta. La justicia de la evaluación se pone en juego, entonces, cuando se efectúa el muestreo de los contenidos que se incluirán en el instrumento a través de las consignas de trabajo (por ejemplo, las preguntas) que éste presenta.

La validez de contenido es también llamada "validez curricular" porque el criterio externo que se emplea para garantizar esta clase de validez es el diseño curricular, ya que en éste se expresan los propósitos y los niveles de aprendizaje, las secuencias y las sugeren-

cias sobre estrategias e incluso, en ocasiones, sobre actividades de aprendizaje y modalidades de evaluación. Con la libertad que debe caracterizar al manejo del diseño curricular por parte del docente, es indudable que este diseño es un recurso fundamental sobre el que se apoyará la estimación acerca de la validez de contenido de un instrumento de evaluación, así como la fuente en la que el docente encontrará la principal inspiración para realizar una selección adecuada y representativa de los contenidos de su curso. Es preciso recordar que cuando se habla de "contenido" se hace referencia a un concepto amplio que comprende todo lo que se quiere enseñar. Se incluyen, por tanto, informaciones, conceptos, principios, procedimientos, valores, actitudes, hábitos, destrezas motrices y todo aquello que se tiene el propósito de que los alumnos aprendan.

En virtud de que lo que se enseña varía de clase en clase, de docente a docente y de región a región, la validez de los instrumentos también es limitada en relación con sus alcances.

Validez predictiva

Esta clase de validez se refiere a la correlación existente entre los resultados obtenidos en una o varias pruebas combinadas y el desempeño posterior del alumno en aquellos aspectos que corresponden al área evaluada por esas pruebas. Los instrumentos de evaluación que se emplean en la escuela deben tener validez predictiva escolar y extraescolar y deben contribuir a pronosticar los éxitos y dificultades de los sujetos del aprendizaje tanto en los aprendizajes posteriores de las asignaturas escolares cuanto en sus desempeños ocupacionales o en su actuación social en general.

Un nivel significativo en la validez predictiva es una condición para un buen instrumento de evaluación. Y esto es así por cuanto no hay aprendizaje escolar cuyo único fin consista en procurar una única ejecución exitosa en un espacio escolar y un tiempo muy limitados. El mismo fin de la educación se define por la voluntad expresa de quien enseña de que los aprendizajes trasciendan el espacio y el tiempo escolares. Es el destino mismo de la educación buscar la perdurabilidad de los aprendizajes. De ahí que, aunque haya algunos instrumentos cuyo propósito específico es la predic-

ción –como es el caso de las pruebas de admisión–, todo instrumento de evaluación empleado con fines pedagógicos tenga, entre otras, la función de predecir. Como antes dijimos, la validez de un instrumento no es absoluta, siempre es limitada. Pero el seguimiento de los alumnos para comprobar la capacidad del instrumento para efectuar predicciones adecuadas es una operación de medición relativamente sencilla, al alcance de todos los docentes que deben encararla como parte constituyente de sus programas de evaluación. Comparar los resultados obtenidos por cada alumno en pruebas diferentes, separadas en el tiempo y en las que el docente asume que hay relación entre los contenidos, es una tarea interesante a la hora de evaluar la enseñanza impartida. Si en una prueba se han detectado errores en las repuestas y se ha brindado atención especial a los alumnos que tenían esas dificultades específicas, el seguimiento de los resultados en una prueba posterior demostrará, a la vez, la validez de la prueba anterior y la adecuación del tratamiento didáctico de los errores que se ha encarado en la acción de enseñanza.

Debemos señalar, sin embargo –aunque esto no exima al docente de hacer el esfuerzo de evaluar la validez predictiva de los instrumentos que utiliza– que los aprendizajes que la escuela promueve tienen plazos más o menos prolongados para su logro pleno. Se han de diferenciar aquellos logros que el alumno puede demostrar en un plazo muy corto, casi contemporáneo con el período de su enseñanza en la escuela, con los aprendizajes que requieren un período algo más prolongado de procesamiento, trabajo y elaboración para una apropiación profunda y aquellos otros aprendizajes que, por el alcance temporal que tiene la maduración personal que requieren, no podrían ser evaluados durante la permanencia del alumno en la institución educativa y ni siquiera en un lapso próximo a su egreso. En estos últimos casos la apreciación de la validez predictiva de los instrumentos se ve limitada por razones técnicas difíciles de superar por el docente y por la institución. Queda en pie, empero, la necesidad de poner atención constante en la capacidad de predicción que tienen los instrumentos de evaluación que se emplean en la educación formal.

Validez de construcción

El programa de evaluación y cada uno de los instrumentos que lo integran deben estar construidos de acuerdo con los principios de la o las teorías didácticas que sostienen el proyecto pedagógico. Debe haber coherencia entre las teorías y las estrategias de enseñanza que se han de emplear o que se han utilizado, las teorías generales y específicas que describen y explican los procesos de aprendizaje que el programa de enseñanza se propone que los alumnos desarrollen y las teorías que fundamentan las elecciones de los tipos de instrumento, la secuencia empleada, la articulación de la enseñanza y de las instancias de evaluación que conforman el programa de evaluación diseñado. Lo mismo ocurre con las normas técnicas que deben respetarse cuando se construye cada tipo de instrumento de evaluación. De la coherencia de las teorías y el buen uso de las normas técnicas depende la validez de construcción del programa y de los instrumentos de evaluación.

Es éste un aspecto que debe ser particularmente cuidado por los docentes porque si la evaluación no tiene un fundamento teórico válido, las decisiones que se tomen a partir de la información recogida por esa vía y de las interpretaciones que se hagan de ella carecerán de legitimidad.

Puede comprenderse, por lo tanto, la importancia de esta clase de validez ya que ella condiciona las otras que antes hemos mencionado.

Validez de convergencia

Esta modalidad de la validez concierne a la relación que existe entre un programa de evaluación o un instrumento y otros programas o instrumentos de validez ya conocida. Cuando un programa o una prueba de evaluación son reemplazados por otros, por ejemplo, debe demostrarse que son, por lo menos, tan buenos o mejores que los empleados anteriormente. La validez de convergencia puede tener gran importancia, pues, cuando se modifican prácticas establecidas en la evaluación para mejorar su validez, su confiabilidad o su practicidad, o cuando se busca mejorarla ampliando el campo de los aspectos evaluados.

Tomemos, por ejemplo, una situación en la que es necesario

encontrar una nueva modalidad de evaluación que permita trabajar adecuadamente con un número muy superior de alumnos al que se acostumbraba atender en la institución. En ese caso, se trata de encontrar un programa e instrumentos más prácticos pero igualmente válidos que los que se empleaban con anterioridad. Los resultados obtenidos con los nuevos programas deberán ser contrastados, en consecuencia, con los obtenidos por medio de la aplicación de los usados antes. La validez de convergencia se apreciará analizando las coincidencias entre ambas series de resultados de la evaluación. Pero si, en cambio, se procura obtener mayor información acerca de un aspecto significativo del aprendizaje, profundizando el examen de los logros obtenidos por los alumnos en competencias o ejecuciones específicas que hasta ese momento se evaluaban de manera más general, en situaciones diferentes o en tipos de expresión distintos, la elección de los nuevos instrumentos deberá ser sometida a un análisis de convergencia para establecer si los resultados de unas y otras series son coincidentes o contradictorios. Si ocurriera esto último, habría que reexaminar tanto los instrumentos empleados previamente como los nuevos, antes de adoptarlos y dar estabilidad al nuevo programa de evaluación.

Validez manifiesta

Esta modalidad de la validez no se refiere a la validez intrínseca del programa o del conjunto de las pruebas o de cada prueba, sino al modo en que ellos se presentan, aparecen, ante el público.

Su significatividad deriva de que los programas y los instrumentos de evaluación deben ser percibidos por los alumnos, sus padres y el público como instancias adecuadas para la evaluación. No pueden ser vistos como absurdas carreras de obstáculos ni como simples o complejos filtros que tienen el único fin de producir una selección, más o menos azarosa y más o menos injusta, entre los alumnos. La razonabilidad de los instrumentos debe ser visible y explicable. Los instrumentos deben ser vistos como formas capaces de dar cuenta de los aprendizajes necesarios de los alumnos, tanto en lo que respecta a la forma como a los contenidos evaluados.

La validez manifiesta exige, entonces, que los instrumentos evalúen de manera verosímil los aprendizajes. Sin esto la evaluación

pierde transparencia y no puede convertirse en autoevaluación. Al perderse la confianza en ella, pierde de este modo su sentido pedagógico y se requiere efectuar las correcciones y operaciones que resulten necesarias para lograr restablecerlo.

Validez de significado

Dietel, Herman y Knuth (1991) añaden a los distintos tipos de validez que se describen habitualmente, la validez de significado, que se mide en respuesta a la siguiente pregunta: ¿las tareas de evaluación tienen significado para los estudiantes y los motivan para alcanzar su mejor rendimiento?

Semejante a la modalidad de validez manifiesta que hemos descrito, la validez de significado se centra, en particular, en la relación que se establece desde la perspectiva de los estudiantes, entre el programa y los instrumentos de evaluación, por un lado, y los procesos de enseñanza y de aprendizaje, por el otro. La evaluación tiene diversos efectos sobre el aprendizaje. Entre los efectos deseados se encuentra el de proveer información al alumno sobre los progresos y los obstáculos que encuentra en el proceso de su aprendizaje y sobre los errores que debe superar, dándole indicios claros acerca de cuál puede ser el origen de sus dificultades. Pero la evaluación también tiene el efecto de incentivar la motivación del alumno por aprender. La superación de dificultades constituye un aliciente para el aprendizaje, así como la reiteración de los fracasos disminuye el interés, si se perciben las dificultades como insuperables. En ese espacio de dimensiones variables para cada alumno, y que se establece entre los problemas demasiado difíciles para ser resueltos y los problemas demasiado fáciles en los que no se puede fallar, se encuentra el punto adecuado en el que debe ubicarse la dificultad que presenta la prueba de evaluación para que tenga significado para el estudiante y constituya un desafío en el que éste ponga su mejor esfuerzo de producción.

Validez de retroacción

La evaluación tiene un efecto de retroacción sobre la enseñanza. Un caso clásico es el que se presenta en las evaluaciones externas, que ejercen una influencia normativa sobre lo que se enseña. Conocidos los contenidos y las técnicas empleadas en las evaluaciones, ellas se convierten para los docentes en los modelos de lo que se debe enseñar y aprender. Lo mismo ocurre, con frecuencia, con los aprendizajes encarados por los alumnos, que se conforman a las pautas establecidas por los docentes en las evaluaciones que administran. Es un efecto, en este caso no deseado, de la evaluación, por cuanto, en vez de ocupar el lugar de un apoyo para el mejoramiento del aprendizaje y de la enseñanza, la evaluación se convierte en la verdadera reguladora de esos procesos. Cuando esto ocurre, el peligro es grande, ya que, como dijimos antes, el dominio de los aprendizajes en la educación es mucho más amplio que el que se puede cubrir mediante la evaluación.

La validez de retroacción, que debe ser medida al igual que los otros tipos de validez, impone la realización de acciones de corrección con el fin de restablecer el verdadero sentido didáctico de los procesos de evaluación.

Confiabilidad

Se dice que un instrumento de evaluación es confiable cuando une exactitud en la medición y sensibilidad para la apreciación de la presencia y las diferencias de magnitud de los rasgos que mide.

Para que los resultados obtenidos mediante la administración de un instrumento de evaluación puedan ser considerados dignos de confianza, deben ser estables, permaneciendo semejantes en todas las ocasiones en que se administre ese instrumento u otro similar. Si, por ejemplo, se administra una prueba o se emplea algún otro tipo de instrumento de evaluación y se obtienen ciertos resultados, el instrumento es confiable si, unas semanas después, administrado nuevamente se obtienen resultados similares. Esa constancia de los resultados demostraría, primero, que el papel del azar es muy pequeño y no distorsiona de modo significativo los

resultados obtenidos. Y, además, que la administración del instrumento ha podido dejar de lado la influencia de factores transitorios que no deberían tener relevancia en la consideración de los resultados del aprendizaje que se quieren evaluar.

Un instrumento confiable permite aislar los aspectos que mide de otros que para el caso se consideren irrelevantes. Así, por ejemplo, una lámpara colgante puede servir para detectar un sismo cercano, pero no puede ser considerada un instrumento confiable pues, al no encontrarse aislada de otros factores, el viento o una mano, por ejemplo, podría ser accionada por otras fuerzas que falsearían la interpretación de su movimiento pendular. De la misma manera, un instrumento utilizado en la escuela con fines de evaluación de los aprendizajes alcanzados por los alumnos debe permitir la consideración de estos aprendizajes separándolos de otros factores que pueden incidir en las respuestas de los alumnos, tales como el azar, la fatiga o los estados de tensión nerviosa de los alumnos o de los docentes.

La confiabilidad depende, también, de la exactitud y precisión con que mide el instrumento. En primer lugar, debe detectar si el aprendizaje o el rasgo que es evaluado está presente o ausente (¿saben o no saben *x*?, ¿tienen *y* destreza o no la tienen?, ¿quieren *z* o no la quieren?, etcétera). Pero debe hacerlo con sensibilidad y sin ambigüedad ante las variaciones de intensidad, profundidad y calidad de las conductas observadas o inferidas.

La confiabilidad de un instrumento depende, pues, de su estabilidad, su exactitud y su sensibilidad. Agregamos a estas condiciones la objetividad, que supone que la evaluación es independiente de la persona que evalúa, esto es, que los resultados reciben una interpretación similar tanto de diferentes docentes como de un mismo docente en distintos momentos.

Entre los factores que afectan la confiabilidad se encuentran los siguientes: la longitud y duración de la prueba, ya que cuando es excesivamente breve su capacidad de cobertura de contenido es escasa y puede estar muy distorsionada por factores de azar. "La muestra limitada de conducta puede dar como resultado una puntuación inestable y poco confiable" (Thorndike y Hagen, 1989). Pero cuando la prueba es excesivamente larga, la duración prolongada puede proporcionar información afectada por la fatiga del alumno. Sin embargo, señala Gronlund (1973) que es preferible el

riesgo de diseñar pruebas más largas para que tengan mayor validez de contenido, siempre que se añadan contenidos diferentes y que no se superpongan con los que ya presenta la prueba. Cuando no se trata de instrumentos para medir velocidad sino de pruebas de competencia o capacidad, el tiempo debe ser suficiente para la realización del trabajo, incluso para estudiantes lentos.

Otro factor que incide sobre la confiabilidad son las condiciones de administración del instrumento. En situaciones puntuales de evaluación formal, la tensión de los alumnos suele ser mayor que en las informales, y la manera en que son tratados por el examinador puede favorecer o perjudicar su desempeño. También tienen influencia sobre los rendimientos las condiciones materiales y físicas de administración de las evaluaciones tales como iluminación, silencio, mobiliario, equipamiento adecuado y comodidad en general de alumnos y docentes.

Alcances de la confiabilidad

La confiabilidad de un instrumento debe permitir cubrir apropiadamente las necesidades planteadas por el programa de evaluación. No todos los instrumentos tienen la misma exactitud y sensibilidad, ya que ella depende de la forma técnica que adoptan y de las características propias que les otorga su construcción específica. Así, en otro orden de mediciones, por ejemplo, un barómetro doméstico puede ser aceptablemente confiable porque las lecturas que permite efectuar brindan la información que, con márgenes de error adecuados a ese uso doméstico, le permiten cumplir con la función a la que está destinado. En una estación meteorológica, en cambio, esas mediciones serían consideradas insuficientemente confiables, pues la sensibilidad y exactitud requeridas para cumplir su función son mucho mayores. Los instrumentos de medición de precisión son los que tienen mayor estabilidad, exactitud, sensibilidad y objetividad. Pero se debe recordar siempre que todas las mediciones no son absolutas, sino que son relativas a los instrumentos de medición que se emplean.

En cada caso, en consecuencia, habrá de determinarse cuáles son los grados de confiabilidad que resultan necesarios para cumplir

con los propósitos que el programa de evaluación asigna a cada instrumento.

No se puede confundir confiabilidad con validez. Es frecuente que los autores de libros sobre este tema afirmen que la confiabilidad es condición necesaria, aunque no suficiente, de la validez (Karmel, 1986; Gronlund, 1973; Popham, 1983; Ebel, 1963; Sachs Adams, 1970; Pidgeon y Yates, 1979). Veamos algunas de estas afirmaciones.

Popham (1983): "Un test carente de (con)fiabilidad no puede ser válido para fin alguno". Ebel (1963): "Para que una prueba sea válida debe ser confiable". Pidgeon y Yates (1979): "Un examen no puede ser válido a menos que sea (con)fiable; pero puede ser (con)fiable sin ser válido". Karmel (1986): "Recuerde que ningún test será válido a menos que sea confiable. Por otra parte, un test puede ser confiable y carecer de validez". Esta exigencia de confiabilidad como condición para hacer posible la validez ha tenido mucha influencia en la teoría y la práctica de la evaluación. Consideramos, sin embargo, que es un principio que debe ser examinado y modificado.

La validez y la confiabilidad de un instrumento de evaluación nunca son absolutas. Hemos visto que para la validez los criterios son los que sirven de pauta y, del mismo modo, que la confiabilidad presenta márgenes de aceptabilidad según el uso que se le quiera dar al instrumento.

Pero por otra parte, en los instrumentos de evaluación de los aprendizajes, cuando se exige confiabilidad como primera condición para lograr cumplir con la exigencia de validez del instrumento, se privilegia un cierto tipo de instrumentos en los que el peso de la subjetividad desaparece. Esos instrumentos de alta confiabilidad son aptos para evaluar un cierto tipo de aprendizajes o rasgos pero no para evaluar todos. Las llamadas "pruebas subjetivas" o "tradicionales" (exámenes o cuestionarios orales, pruebas escritas de ensayo o disertación o de respuesta extensa; pruebas de respuesta restringida o de cuestionario; monografías, tesinas y tesis; evaluación de portafolio; estudios de casos, entre otras) están destinadas a evaluar la producción del alumno sin pautas totalmente estructuradas ya que le dan libertad para escoger la información, organizarla, expresarla y presentarla. En estos casos la validez aumenta a expensas de la confiabilidad. Esta última no es, entonces, condición

necesaria de la validez sino, por el contrario, una limitación de la validez en ciertas circunstancias, aquellas en las que se busca que la respuesta del alumno sea original y permita demostrar competencias que no pueden ser evaluadas por medio de instrumentos cuyas respuestas están previamente estructuradas por los evaluadores.

Conviene señalar, sin embargo, que según se trate de propósitos de alcance e impacto muy importante en las vidas de las personas (exámenes de admisión o de selección, por ejemplo) o de las instituciones, la transparencia de las situaciones de evaluación constituye una exigencia ineludible para la evaluación y que, para ello, se requiere una alta confiabilidad además de validez predictiva.

La confiabilidad es, por lo tanto, una característica importante, difícil de aplicar adecuadamente y que exige un manejo técnico muy correcto de las técnicas de la evaluación puestas al servicio de los propósitos que ésta tiene asignados en los procesos de enseñanza.

Practicidad

Además de las características mencionadas –validez y confiabilidad– hay que considerar una tercera, de mucha importancia en los programas de evaluación.

La practicidad de un programa o un instrumento resulta de la conjunción de tres aspectos: su administrabilidad, la facilidad de análisis e interpretación de sus resultados y elaboración de conclusiones y, por último, de la evaluación de la economía de tiempo, esfuerzo y costo de su utilización.

Administrabilidad

La administrabilidad de un programa o un instrumento es un producto de varias características:

1. el tiempo de trabajo de los docentes que insume su diseño y construcción;
2. el tiempo que exige su puesta en práctica;
3. la claridad de la/las consigna/s, es decir, de las directivas que se dan a los estudiantes;

4. los materiales, equipos y lugares especiales necesarios para su administración;
5. la cantidad y preparación de las personas indispensables para la administración, cómputo, análisis e interpretación de los resultados y elaboración de las evaluaciones.

Cuanto menor sea el tiempo necesario para el diseño y la elaboración de las conclusiones y menos especializado el personal necesario para la puesta en práctica y el cómputo, mayor será la administrabilidad del programa o instrumento.

Análisis e interpretación de los resultados y elaboración de conclusiones

Se puede advertir que hay una correlación inversa entre el tiempo que demanda la construcción de un instrumento y el que exige el análisis e interpretación de sus resultados. Es necesario señalar, sin embargo, que el diseño de programas e instrumentos de evaluación siempre requiere mucho tiempo, también cuando se trata de pruebas "subjetivas". El tiempo de "corrección" de las pruebas o de registro de observaciones suele ser prolongado pero la tarea puede ser facilitada por la construcción de claves o grillas de corrección elaboradas previamente a la administración de la prueba, sea escrita, oral o de evaluación de procesos o productos a través de la observación. Ellas permiten disminuir el tiempo y los problemas que se les plantean a los docentes en el momento de la corrección, les dan mayor seguridad y contribuyen a mejorar la confiabilidad de los resultados de la evaluación. Se deben emplear con precaución para no destruir el espíritu de estas pruebas que es la libertad de presentación y organización que se le brinda al alumno. De otro modo sería preferible utilizar instrumentos de tipo objetivo, esto es, instrumentos que admiten una y sólo una respuesta correcta y no diferentes calidades de respuesta que son más difíciles de analizar.

Economía

Es importante, como hemos dicho antes, que el programa de evaluación sea económico en tiempo, trabajo y costo. Es conveniente

examinar previamente estas características para determinar la conveniencia de su empleo.

Integración de los tres aspectos

Los tres aspectos relativos a la practicidad del programa o de cada instrumento deben considerarse conjuntamente para lograr una evaluación integral de este rasgo.

Cada aspecto presenta ventajas y desventajas que se compensan o equilibran cuando se analizan los otros dos. Un costo relativamente elevado, por ejemplo, podría ser compensado por un alto grado de administrabilidad, o un tiempo prolongado dedicado a la construcción de un instrumento puede ser equilibrado por un tiempo menor destinado a la corrección e interpretación de los resultados. Del mismo modo, un tiempo extenso de construcción puede compensarse por la facilidad de administración simultánea a un grupo muy numeroso de alumnos.

La practicidad no es, seguramente, la característica más importante de un programa o de un instrumento de evaluación. Pero sí adquiere relevancia cuando de ella depende la frecuencia de la evaluación. Se observa que, por lo general, cuando el programa es práctico o cuando el instrumento es de fácil administración e interpretación se evalúa más frecuentemente y ello contribuye a estructurar mejor los procesos de enseñanza y los procesos de aprendizaje de los alumnos por parte del docente y de los mismos alumnos.

Utilidad

Íntimamente asociada a las otras tres características que mencionamos –validez, confiabilidad y practicidad–, la utilidad de una prueba resulta de su capacidad para satisfacer las necesidades específicas relacionadas con los procesos de enseñanza y aprendizaje. Si los resultados de la evaluación no son útiles para la orientación de los alumnos, del docente, de la escuela como institución, de los padres de los alumnos, de los sectores interesados en la calidad de la educación, poco importa que, en un marco restringido, el programa y cada instrumento sean confiables y prácticos. En lo que

respecta a la validez, ya hemos dicho que es relativa a los criterios y, tratándose de públicos diferentes, dichos criterios pueden ser heterogéneos. Para que la evaluación sea útil, la validez para alguno o algunos de esos públicos es indispensable.

La pregunta esencial en el diseño de un programa de evaluación es, por lo tanto, ¿qué usos se les dará a los resultados obtenidos? Son los implicados en los procesos de evaluación quienes deben dar las respuestas.

Bibliografía

Cardinet, Jean: "Evaluer sans juger", *Révue Française de Pédagogie*, nº 88, julio-agosto-setiembre de 1989.

Dietel, R. J., Herman, J. L. y Knuth, R. A.: "What does Research say about Assessment?", NCREL, Oak Brook, 1991.

Ebel, Robert L.: "Must all Tests be Valid?", en *Human Learning in School*, ed. John De Cecco, Nueva York, Holt, Rinehart & Winston, 1963.

——: *Fundamentos de la evaluación educacional*, Buenos Aires, Guadalupe, 1977.

Garrett, Henry E.: *Estadística en psicología y educación*, Buenos Aires, Paidós, 1966.

Gronlund, Norman E.: *Medición y evaluación en la enseñanza*, México, Pax, 1973.

House, E. R., *Evaluación, ética y poder*, Madrid, Morata, 1994.

Karmel, Louis J.: *Medición y evaluación escolar*, México, Trillas, 1986.

Perrenoud, P.: *La construcción del éxito y el fracaso escolar*, Madrid, Paideia-Morata, 1990.

Pidgeon, Douglas y Yates, Alfred: *Evaluación y medida del rendimiento escolar*, Madrid, Anaya, 1979.

Popham, James W.: *Evaluación basada en criterios*, Madrid, Magisterio Español, 1983.

Sachs Adams, Georgia: *Medición y evaluación*, Barcelona, Herder, 1970.

Thorndike, Robert L. y Hagen, Elizabeth P.: *Medición y evaluación en psicología y educación*, México, Trillas, 1989.

LA EVALUACIÓN DE LAS PRÁCTICAS DOCENTES Y LA AUTOEVALUACIÓN

María del Carmen Palou de Maté[1]

A Pelusa
y nuestros hijos

Este escrito pretende explicitar algunos conceptos que permitan al docente repensar la evaluación y a la vez brindar líneas de trabajo que contribuyan a mejorar su práctica.

En una primera instancia se intenta dar cuenta de los diferentes modos que asume la evaluación en el ámbito de la didáctica, entendiendo que su conformación actual responde tanto a la construcción histórica del concepto mismo como a los entrecruzamientos contextuales que recibe. Esto permite analizar la evaluación desde una dimensión política, desde una perspectiva psicológica y desde un marco sociológico, así como también a partir de la forma en la que ha penetrado en ella el cientificismo.

Estas dimensiones que se entrecruzan en la práctica, determinan por sus fuerzas una configuración particular de la evaluación, generando muchas veces contradicciones que pueden explicitarse cuando se intenta abordar algunos aspectos teóricos desde la práctica y para la práctica.

En una segunda instancia se propone tomar la tarea del docente en el aula como un proyecto, al cual la evaluación le permite analizar su propuesta didáctica en vinculación con otros proyectos que constituyen su contexto, resaltando la implicancia de éstos en su tarea.

Desde este lugar se plantea la autoevaluación de la enseñanza,

1. Profesora asociada regular, Área Didáctica, Facultad de Ciencias de la Educación, Universidad Nacional del Comahue.

entendiendo que ofrece genuinas posibilidades de revisar las prácticas y la autoevaluación del alumno como un modo de trabajo que facilita la vinculación de éste con el conocimiento.

El lugar de la evaluación

Al iniciar una reflexión enmarcada en el ámbito de la didáctica, es importante definirla para encontrar el sentido de las dimensiones que se analicen.

Se concibe a la didáctica como la teoría de la enseñanza que tiene como finalidad el estudio de la intervención docente en los procesos de construcción del conocimiento por parte de los alumnos en un contexto sociohistórico determinado.

Este punto de partida requiere que se reconozca el fenómeno educativo como un fenómeno complejo, donde tal complejidad obliga a dar cuenta de los múltiples entrecruzamientos que lo constituyen, no sólo en un aquí y ahora, sino atendiendo fundamentalmente al proceso histórico a partir del cual se construye una práctica o suceso educativo.

Es así como la evaluación en el aula ocupa un lugar importante, ya que es constitutiva de las prácticas pedagógicas mediando el encuentro entre el proceso de enseñar y el de aprender.

Marco sociohistórico

Haciendo un rastreo del concepto de evaluación, pueden encontrarse en los grandes pedagogos distintas concepciones de él, pero el que deviene de la Revolución Industrial es el que aparece con más fuerza, aunque muchas veces no de forma explícita.

Desde esta perspectiva la evaluación se articula entre el sistema productivo y el sistema escolar, relación ésta que se manifiesta también en otros ámbitos del campo pedagógico, donde la demanda industrial se va reconvirtiendo en términos educativos en todo el quehacer docente, como por ejemplo la concepción de número, el vocabulario, la actitud frente al adulto.

Analizada la evaluación desde esta perspectiva a partir de su instrumentación didáctica, los objetivos comportamentales tienen

una función central, ya que, explicitados antes de iniciar el proceso de enseñanza y aprendizaje, serán el referente para confrontar los resultados de este proceso. He aquí la función de la evaluación: verificar los resultados obtenidos en función de objetivos prefijados, metas cuidadosamente especificadas con anterioridad.

Esta evaluación que se relaciona con el proceso de industrialización de Estados Unidos se refiere especialmente al desarrollo de los conceptos vinculados al "manejo científico del trabajo" de Frederick Taylor. Es decir que, desde el surgimiento de la significación de este vocablo, la connotación ideológica está ligada al ámbito del control administrativo, esto es, tiene un carácter técnico que con métodos e instrumentos intenta dar cuenta y rendir cuenta.

El principio de actuación que implica planificar, realizar y evaluar pasó de la empresa a los centros educativos y se constituyó en pautas para el desarrollo de las tareas de índole pedagógico-didáctica. La segmentación técnica del trabajo tuvo su reflejo en la segmentación de la actividad docente. Allí aparecen los especialistas en currículo, en planificación, en evaluación, llegando, en este último caso, a la objetivación del número como garantía de objetividad y rigor científico.

Esto nos lleva a pensar que no es casual que cuando el docente utiliza la palabra "evaluación" lo haga con la mira puesta en los resultados de aprendizaje previstos curricularmente, es decir, refiriéndose a la acreditación (planificar-implementar y evaluar), ya que ésta es un punto que articula la escuela con los requerimientos sociales especialmente ligados al mundo del trabajo.

Analizando esta analogía ("evaluación" como sinónimo de "acreditación"), se comienza a percibir la dificultad que ocasiona la polisemia del término.

Desde aquí es interesante resaltar los múltiples significados que tiene la palabra "evaluación". Bertoni, Poggi y Teobaldo (1995) abordan con claridad esta problemática señalando lo siguiente:

> En primer lugar, evaluar parece involucrar el hecho de medir con precisión; en este sentido, expresar una cantidad precisa, cifrada. En otras palabras, alude a expresar una medida cuantificada.
>
> En segundo lugar, evaluar implica operaciones como estimar, apreciar, aprehender, en otros términos, pronunciar un juicio cualitativo y, eventualmente, aproximativo sobre una realidad.

Por su parte, Ardoino (1988) sostiene que es tanto "la marcha crítica de la experiencia cotidiana como el instrumento más sofisticado de estimación o apreciación de proyectos más organizados...".

Estas consideraciones aportan elementos para plantear que "evaluación" es un término polisémico[2] que tiene su origen en procesos sociales e históricos particulares. En este sentido el término "evaluación" no tiene identidad disciplinaria; se emplea en campos tan disímiles como la física nuclear, la bioquímica, la economía, la psicología y la propia educación. Todos lo emplean indistintamente sin lograr una conceptualización rigurosa en su propio campo.

Si bien se reconocen los orígenes y el uso contemporáneo de esta palabra, de los que no es posible desprenderse en los análisis interpretativos, es necesario desentrañar su significación en las prácticas evaluativas que acontecen en el salón de clase.

Dimensión política e ideológica

Deslindando el uso de este término en la vida cotidiana y acotándolo al ámbito educativo, en su utilización es fácil ver que "...se busca una imagen de seriedad académica, conceptual y administrativa" (De Alba y otros, 1988), de la que no está exento el docente.

En la necesidad de "evaluar bien" subyace un gran temor a equivocarse y, en este sentido, es la ciencia la única que puede ayudar a dar respuesta a este problema.

El valor que la sociedad da a la evaluación, generalizada en el ejercicio de muchas profesiones, proporciona al docente un buen soporte a su tarea. Popkewitz (1986) marca con fuerza esta idea diciendo que "la capacidad de la ciencia como legitimadora de la actuación profesional es una de las causas del desarrollo de la evaluación...".

Es frecuente que detrás de la demanda de evaluación se pueda "ver" la necesidad de dar un marco de cientificidad a la tarea que se realiza, pero si se profundiza lo que se percibe es un mejor instru-

2. "...los signos lingüísticos a partir de los cuales se construye la palabra evaluación no generan un significado del concepto; no hay univocidad en lo que se entiende por evaluación, de hecho tenemos una palabra que genera una multitud de nociones" (De Alba y otros, 1988).

mento para evaluar, un instrumento que clarifique, que proporcione datos sobre los resultados obtenidos por el alumno de la forma más objetiva posible.

Aquí aparece la primera contradicción, que es la búsqueda de neutralidad en una actividad esencialmente valorativa.

Desde este lugar se concibe a la evaluación como una actividad neutra, que desde su propia cientificidad (que va más allá de las personas) legitima la acción pedagógica, pero a su vez al definirla se la concibe como una fuente de valor. Popkewitz (1986) afirma que en el seno de la evaluación

> se produce una contradicción entre su función de juzgar el valor y su condición de actividad tecnológica neutral, contradicción que responde a una cierta postura ideológica.

A esto Gimeno Sacristán (1985) aporta:

> Si toda actividad humana se encuentra penetrada por una inevitable dimensión intencional, ética y teleológica, la esencia misma de toda actividad evaluadora es su componente axiológico [...] es un componente inseparable del mismo objeto de su actividad.

La evaluación, por ser una práctica humana, posee también en sí una dimensión política, ligada al ámbito del poder.

No es casual que las demandas de evaluación partan de los organismos, en este caso educativos, con estructuras de poder consolidadas.

Es la información la que lleva en sí misma una cuota de poder y permite un mejor control, dando lugar a la realización de estrategias de acción coherentes a perspectivas ideológicas particulares.

Bajo este enfoque la evaluación puede concentrar o distribuir poder acorde a quienes sean los destinatarios de la información y el uso que de ella se haga.

En el aula esto se ve con claridad en el contrato pedagógico, donde la explicitación de las normas y usos que se realiza de los datos de la evaluación concentran o distribuyen poder en el docente o en los alumnos.

La evaluación en el aula

Acerca del concepto de evaluación

Tanto la idea de cientificidad como el uso del poder, se incorporan a la tarea del docente, permeando su actividad evaluativa. La palabra evaluación en el ámbito escolar está casi homologada a la idea de acreditación,[3] esto lleva a que en el aula se cruce con la idea de nota, de calificación. Es una palabra que está especialmente significada en el discurso pedagógico,[4] y fuertemente significada desde el control y desde el poder que tiene el docente.

Evaluar es una acción que se realiza con mucha frecuencia en la vida cotidiana, tomando criterios que provienen del sentido común.

Para incorporar este concepto a las prácticas educativas, es conveniente darle rigurosidad al tratamiento del mismo, lo que implica señalar los elementos que se consideran constitutivos para dar cuenta de ella.

En términos genéricos evaluar es valorar, lo que lleva en sí emitir un juicio de valor acorde a marcos axiológicos,[5] tendiente a la acción.[6] Esto implica aproximarse al "objeto"[7] tomando como punto de partida la descripción,[8] comprensión y explicación del mismo.

3. "Acreditación" se define como la certificación de conocimientos curricularmente previstos. Responde a una demanda social e interinstitucional.

4. Resulta muy interesante el análisis de Basil Bernstein (1990), en relación con el discurso pedagógico como discurso instructivo y la inclusión de éste como discurso regulativo, donde la evaluación ocupa un lugar importante.

5. Referido a un sistema de valores que surgen de diferentes visiones del mundo y que pueden darse en forma tácita o explícita.

6. Estableciendo la diferencia entre comportamiento como respuesta observable y actuar en el sentido de acción informada con conciencia de la relación medio-fines, superando la visión instrumentalista que ha tenido hasta el momento.

7. Tomado en un sentido epistémico. En general hace referencia a procesos, fenómenos, comportamientos, valores –en términos económicos–, proyectos individuales o colectivos, prácticas sociales, servicios o productos comerciales, políticos, estratégicos, planes y programas, personas.

8. Cabe aclarar que en la descripción ya se plantea una valoración que puede verse en el reconocimiento de aquellos datos que proporcionan elementos al juicio y el dejar de lado otros que parecieran no tener suficiente peso para ello. Se parte de que toda actividad humana es valorativa pero tiene como sentido en este caso sólo la aproximación al objeto, y en este trabajo se lo concibe orientado a la acción.

Al hablar de evaluación, las dimensiones, características y complejidad del objeto para evaluar señalan las primeras aproximaciones a él. Así, por ejemplo, si se trata de un proyecto social (aunque sea de carácter pedagógico), reviste características diferentes a las que tiene el fenómeno del aprendizaje.

Este señalamiento marca los primeros límites de las extrapolaciones metodológicas que algunas veces se realizan en el campo de la evaluación, por esto difícilmente se pueda tomar la propuesta metodológica de la evaluación "iluminativa"[9] de Parlett y Hamilton para dar cuenta de los aprendizajes de los alumnos, aunque dicha palabra tenga una buena resonancia en el campo de la didáctica.

Tratando de profundizar en el sentido de la evaluación de los aprendizajes, ésta se puede concebir de dos maneras, como inherente a la dinámica interna del enseñar y del aprender[10] en el reconocimiento mutuo de ambos procesos, y como acreditación, que implica dar cuenta o rendir cuenta de los resultados de aprendizajes logrados en un tiempo y nivel de escolaridad determinados.

Este último responde a una demanda social que requiere de la escuela la certificación y legitimación de conocimientos: es la convalidación de un mínimo de aprendizajes curricularmente previstos, es la certificación de que se han logrado ciertos productos o resultados de aprendizajes planteados en programas y planes de estudio.

La acreditación certifica que se ha alcanzado determinado nivel dentro de lo que Perrenoud (1990) llama "jerarquías de excelencia".[11] Las formas en que se gestan los juicios de excelencia tienen que ver con la organización de la escolaridad, con los contenidos de la cultura escolar y con las prácticas pedagógicas.

Retomando la idea de evaluación como inherente al proceso de enseñar y de aprender, se puede afirmar que una de las razones fundamentales que la justifica es la de concebir al conocimiento

9. La evaluación ha de abarcar no sólo los resultados de la enseñanza sino a ésta en su totalidad: fundamentación, desarrollo, dificultades.

10. Generalmente basados en la psicología y, a partir de ello, en tres grandes perspectivas: conductista, cognitiva y sociocultural.

11. Para el autor, excelencia es "la imagen ideal de una práctica dominada a la perfección, cumplida, auténtica".

como un proceso de construcción compartido entre el que enseña y el que aprende, donde el intercambio de significados se produce en el reconocimiento de la naturaleza del aprendizaje de los alumnos y en la posibilidad de replantear las estrategias de enseñanza. En este sentido, es una evaluación orientada a una situación de interacción.

En síntesis, interesa resaltar la naturaleza diferente de ambos conceptos, reconociendo que en la complejidad de la práctica educativa ambos se entrecruzan, se encuentran y/o se homologan.

Es decir que la evaluación de los aprendizajes puede cumplir con varias finalidades: la primera –y más importante para el campo de la didáctica– es proporcionar datos que permitan desplegar diferentes estrategias de enseñanza. El partir del reconocimiento y la comprensión de las formas de aprender, en cuanto a su estructuración y producción –sin caer en un isomorfismo–, permite replantear las estrategias de enseñanza.

La segunda, acreditar, es la certificación de conocimientos curricularmente previstos, es la certificación de que se han logrado determinados productos planteados en planes o programas de estudio. Este requerimiento social e interinstitucional es el que permite la movilidad de los alumnos en las instituciones y a la vez los inserta en el campo laboral.

La acreditación y los vínculos con el conocimiento

Por la incidencia que tienen en el campo social, la acreditación se sobredimensiona, ocupando un lugar privilegiado a la vez que mediatiza la relación del alumno con el conocimiento.

En este sentido son muchos los planteos que se hacen en torno del examen –tomando a éste como el prototipo de la instancia de validación para la acreditación–, entre los que merecen señalarse los de Gimeno Sacristán (1985) y Elliot (1990).

Sacristán (1985) afirma que

> la evaluación se convierte en un elemento inicial que condiciona todo lo demás [...] su fuerza determinante es tal, gracias a la introyección psicológica que el sujeto hace de sus consecuencias escolares y sociales, que define por sí misma el método que se usa [...] Al fin y al cabo lo que

en el proceso de enseñanza-aprendizaje se haga será valorado o no en el momento de evaluar y, seguramente poco a poco, el alumno hace lo que tiene cabida en el momento de ser examinado.

Un valioso aporte para comprender estas prácticas lo realiza Elliot (1990) cuando en su libro *La investigación acción en educación* relata la investigación T.I.Q.L., Teacher-Pupil Interaction and the Quality of Learning, llevada a cabo en Inglaterra en 1979.

El planteo central se basa en la siguiente cuestión:

> en qué medida es posible enseñar para la comprensión en un sistema educativo que se rige en último término por las exigencias de los exámenes.
>
> Aunque han estado muy ocupados escribiendo durante la clase, han aprendido poco, sus esfuerzos se centran en copiar la información en sus cuadernos, para aprenderla más tarde [...] se produce un fenómeno de posposición de los aprendizajes [...] Más tarde significa el estudio previo a los exámenes. La tarea diaria cierra el proceso de anotación. El período de aprendizaje es anterior al examen.
>
> Los exámenes son antieducativos porque reemplazan la motivación intrínseca de aprender para la vida, por la motivación para aprender al corto plazo con el único objetivo de aprobar los exámenes.

A continuación se plantean algunos aspectos relevantes de la acreditación, que obligan a considerarla seriamente y tienen que ver con las diferentes formas de relación del alumno con el conocimiento, anteriormente desarrollados.

- *En el momento de ejecución de las pruebas por parte de los alumnos*

Las palabras de Landsheere (1978) son claras cuando afirma que en la situación escolar habitual el maestro conoce a cada uno de sus alumnos y, en consecuencia, está en condiciones de dosificar, matizar sus notas, en función de un efecto deseado: en un caso, alienta mediante la estimación exagerada de un trabajo; en otro, muestra una severidad excepcional para causar un impacto del cual se espera un efecto saludable. En estas situaciones el maestro procede deliberadamente, con plena conciencia.

Algo muy frecuente ocurre con los fenómenos de estereotipia y de halo.

a) *Estereotipia.* Se entiende por estereotipia una inmunidad más o menos manifiesta en cuanto al concepto que se tiene del alumno. La estereotipia resulta de una contaminación de los resultados. Un primer trabajo mediocre hace pensar que el segundo también lo será. Si esto sucede, la tendencia a otorgar una nota mediocre al tercer trabajo aumenta aún más, y así sucesivamente...

Sería un error creer que la estereotipia influye solamente en las evaluaciones netamente subjetivas...

Alcanza a ejercicios tan "objetivos" como un dictado de ortografía. Ejemplifica el autor que algunos padres, a veces muy competentes en la materia que tenían de tarea sus hijos en el hogar y en la que incidentalmente ayudaron, no sintieron cierta desilusión o asombro al enterarse de "su" nota.

Sin embargo sería injusto deducir que todos los maestros acusados de estereotipia carecen de sentido de responsabilidad. Son los métodos de evaluación los que deben ponerse en tela de juicio, y no los hombres.

b) *Efecto de halo.* El efecto de halo presenta un marcado carácter afectivo. A menudo se sobreestiman las respuestas de un alumno de buen aspecto, con mirada franca y acción agradable. Sin embargo, no es generalizable; existen profesores que por anticonformismo favorecen cierto desaliño u originalidades.

• *Luego de la corrección de las pruebas por parte de los docentes*

Después de un período de tiempo más o menos largo, a través de las pruebas de evaluación o el comentario de otro colega, el maestro tiene "idea" de cómo es un alumno.

Esa idea (estereotipia) hace que el docente espere de él ciertos resultados, prejuicio que actuará en el alumno en forma positiva o negativa sobre su futuro rendimiento. Se dice que un alumno es en la medida de lo que se espera de él.

Es común encontrar docentes que, con muy buena disposición y depositando la confianza en la capacidad intelectual de un alumno, le devuelven su trabajo diciendo "Tú eres capaz de mucho más" mientras que a otros les dicen: "Está bien, yo sé que es lo mejor que puedes hacer". Hay expectativas y exigencias de desafíos cognitivos diferentes en ambos casos.

Evaluar para enseñar y evaluar para acreditar

En síntesis, se podría afirmar que el peso otorgado socialmente a la acreditación, sumado a la polisemia del término "evaluación", dificulta enormemente la tarea de darle a ésta el lugar privilegiado que se merece dentro del campo de la didáctica.

Cabe resaltar que si un docente se acerca a sus alumnos tratando de comprender la lógica de apropiación de los conocimientos, la forma en la que van construyendo sus saberes, entabla con ellos una relación cualitativamente diferente de la de otro que asume una permanente actitud verificadora de los conocimientos adquiridos por los alumnos. Esta doble mirada del aprendizaje de los alumnos, permite un análisis particular de la intervención docente, reubicando la evaluación en el campo de la enseñanza.

Es innegable reconocer el valor de la evaluación que centra la mirada en la comprensión de los procesos de aprendizaje articulando desde allí su propuesta de enseñanza, pero esto no implica un menosprecio por la acreditación, y ni siquiera pensar que una propuesta excluye la otra, ya que el hacerlo implicaría desconocer que la enseñanza es una práctica social y que como tal le corresponde la legitimación de conocimientos. Pero sí interesa darle a esta dimensión de la práctica que llamamos acreditación el lugar que le corresponde, sin sobredimensiones que obturan genuinas prácticas pedagógicas.

Evaluar para organizar el enseñar y evaluar para acreditar se integran en la práctica, pero no debieran confundirse; ambos constituyen la práctica pero no son de la misma naturaleza, ya que responden a finalidades diferentes teniendo sistemas referenciales diferentes.

Acerca del concepto de proyecto y su relación con la evaluación

El concepto de proyecto al que hace referencia el presente escrito es el que plantea Litwin (1992) como:

> [...] propuesta de acción de carácter experimental. Consolida acciones encuadradas en una meta a alcanzar susceptible de ser modificada.

> Un proyecto responde o esconde, según los casos, una concepción del hombre y del mundo, está determinado por una propuesta política, en términos de proyección social y en los proyectos educativos, contiene además una concepción de enseñanza y de aprendizaje...

También Zemelman (1987) incorpora valiosos aportes a la conceptualización de proyecto, tanto cuando dimensiona al sujeto que lo incluye, afirmando que

> La idea de proyecto supone la existencia de un sujeto capaz de definir un futuro como opción objetivamente posible, y no como mera proyección arbitraria,

como cuando lo significa socialmente:

> Es gracias a los proyectos que los sujetos establecen una relación con la realidad, y se apoyan en su capacidad de transformar esa realidad, en cuanto ésta es contenido de una realidad social.

Por su parte Ardoino (1993) marca un posicionamiento del sujeto cuando afirma que

> el proyecto como una intención filosófica o política (es) una intención orientada por valores en búsqueda de realización. Abordar las prácticas de la enseñanza desde este lugar requiere superar el paradigma racional de la explicación y moverse en el terreno de la implicación. Significa trabajar sobre las representaciones, los procesos de influencia, las prácticas sociales. Inventar nuevas formas de regulación a través de dispositivos adecuados.

Se puede pensar en un proyecto curricular o en un proyecto institucional, pero en este trabajo se hará referencia al proyecto pedagógico-didáctico que se da en la clase y que llevan adelante básicamente el docente y sus alumnos, los que, al igual que el resto de los proyectos sociales, se entrecruzan y enmarcan mutuamente.

No se puede desconocer que la institución escolar está cruzada por un "tono" evaluativo que impregna toda su actividad. Allí las acciones realizadas por sus actores son valorizadas –por ellos y por otros– en forma inmediata, lo que dificulta una objetivación de la realidad que permita su comprensión.

La conceptualización que ya se presentó del término "evaluación" puede ayudar a construir nuevos modos de trabajo que superen la función clásica de control, para integrarla al campo de la comprensión.

Evaluar es valorar, otorgar valor desde referentes[12] axiológicos; en este caso, el que subyace al proyecto y que va a marcar las líneas de estrategia metodológica para recolectar información y analizarla, así como también las acciones posteriores a la valoración.

Evaluar implica describir, analizar, interpretar y explicar, lo que permite comprender la naturaleza del objeto y emitir un juicio de valor, el que está siempre orientado a la acción.

Realizar una evaluación desde esta perspectiva supone:

- Apertura conceptual para dar cabida en la evaluación a resultados no previstos y acontecimientos imprevisibles.
- Apertura de enfoques para dar cabida a la recolección de datos, tanto sobre procesos como sobre productos.
- Apertura metodológica. La primitiva e inflexible estrategia formal cambia para incluir procedimientos informales. Del monismo al pluralismo metodológico.
- Apertura ético-política. La evaluación proporciona información a todos los participantes y recoge opiniones e interpretaciones de todos los grupos de intereses implicados en un proyecto educativo... (Gimeno Sacristán, 1985).

La evaluación como se plantea en este trabajo no busca verdades universales y "objetivas" sino más bien intenta ofrecer a los destinatarios que constituyen una comunidad de mentes, argumentos que lleven a un accionar adecuadamente fundado. Se trata de una tarea argumentativa, contextuada y dirigida a la acción.

Las palabras de Ernest House (1994) muestran con claridad lo expresado:

> Probar algo supone satisfacer el conocimiento de una audiencia universal con respecto a la verdad, más allá de toda duda. Producir una prueba aceptable para una audiencia universal que comprenda a todos los hombres racionales exige superar las particularidades locales o históricas [...]. Si la evaluación se limita al conocimiento cierto produci-

12. Formas de legitimación, muchas veces ligados a un sistema de normas.

do por un razonamiento deductivo o inductivo estricto, tiene que dejar de lado gran parte de la capacidad de razonamiento que utilizan las personas en forma habitual para regir sus vidas. Esta limitación es el resultado de confundir la racionalidad con la lógica. No son idénticas...

La evaluación del proyecto pedagógico-didáctico

La propuesta de este trabajo es encontrar en la práctica profesional del docente un proyecto en el cual, como en todo proyecto, la meta[13] direccione la acción, y sea esta direccionalidad lo que le otorgue sentido a su tarea.

Esta mirada en la meta, cuando puede ser entendida en el contexto sociohistórico[14] en el que se inscribe el proyecto, le permite al docente, a través de la evaluación y en una permanente confrontación, interpretación e interpelación de la relación entre estrategias y direccionalidad, ir superando las razones técnicas[15] de su práctica.

El proyecto pedagógico-didáctico así concebido se constituye a partir de los entrecruzamientos contextuales de su inserción en la comunidad; evaluar los aprendizajes y evaluar la enseñanza se dimensionan y significan en función del proyecto en el que se encuentran.

Este proyecto se plantea como un proceso variado que recibe determinaciones desde diferentes instancias, y al analizar desde allí el fenómeno educativo supone múltiples lecturas y obliga también a utilizar diferentes metodologías para abordarlo. A esto le cabe una advertencia, ya que si con fines prácticos se usan instrumentos simplificadores y atomizadores de esa realidad se impide aprehenderla en su complejidad, y al fragmentarla se desnaturaliza.

13. Tiene un sentido de finalidad, orientadora de la acción, superando la propuesta del enfoque sistémico que define la meta como la distancia mensurable entre el estado actual de un proyecto y al que se llega con el logro de objetivos fijados con anterioridad.

14. Hoy más que nunca las políticas educativas son claras en la Argentina, y se pueden significar a través de las propuestas de reforma enmarcadas en la Ley Federal de Educación.

15. La razón técnica está solamente ligada al control de la relación medios-fines sujeto a reglas conocidas sin poder seguir –en un razonamiento práctico y basado en la interpretación– un accionar que le permita actuar con libertad y autonomía.

Esta perspectiva, que posibilita diferentes miradas de lo que acontece en el aula, permite a la evaluación ocupar un papel mediador, ya que es ella la que recoge información sobre las repercusiones que tiene el proyecto en su contexto y los impactos que de él recibe. Esto impide tanto dejar de lado aspectos que habitualmente no son tenidos en cuenta y que inciden en la práctica, como sobredimensionar otros que la penetran, distorsionándola. Así por ejemplo la dimensión política, que durante años estuvo ausente en los análisis, hoy pasa a ser la única causa que configura la realidad educativa actual; o la valoración de los aprendizajes de los alumnos como respuesta inmediata a la enseñanza en una relación biunívoca; o el desconocimiento del contexto institucional y su fuerte incidencia en la práctica docente.

Estas dimensiones, que pueden estar entrecruzadas con otras, encuentran su lugar si la evaluación puede proporcionar información desde diferentes lugares que constituyen el proyecto, ayudando a evitar recortes parciales simplificadores que distorsionan e impiden mejorar la comprensión del fenómeno educativo en su complejidad.

El valor de pensar la práctica del docente como un proyecto e incluir en él su evaluación permite visualizar esta complejidad. En este sentido Edelstein (1995) ubica la evaluación cuando sugiere

> poner en el centro el proyecto educativo o propuesta de enseñanza concebida holísticamente como punto de referencia. De allí asignar a la evaluación la misión de identificar las concreciones y niveles de desarrollo, reconociendo incluso las tensiones y contradicciones que se generan en el entramado de su realización.

Este trabajo se propone tener una visión de la realidad que recupere las múltiples dimensiones que constituyen la práctica del docente en el aula, y a partir de allí ir integrando y resignificando en un todo –complejo y a veces contradictorio– distintas temáticas o problemáticas que se le plantean diariamente.

¿Cómo iniciar la tarea de evaluación del proyecto?

Se considera como punto de partida que la evaluación está ligada a valores y que supone –aunque no siempre explicitada– una

concepción del hombre y del mundo. Desde estos marcos tan generales hasta la instrumentalización, es importante poder encontrar propuestas metodológicas que articulen ambos polos "intentando" no perder la visión de totalidad y complejidad del objeto por evaluar.

Las diferencias terminológicas, aunque parezcan complicadas en una primera lectura, pueden ayudar a comprender un sentido de la evaluación diferente, las que se irán trabajando con ejemplos más adelante.

Cabría distinguir:

- *el referente* es el elemento exterior al que el "objeto" que se ha de evaluar puede ser relacionado, generalmente ligado a un modelo;
- *el referido* es el "objeto" mismo que será evaluado;
- *el sistema de referencias* tiene que ver con un campo científico que pueda dar cuenta del "objeto" por evaluar; proporciona una óptica, un tipo de análisis, un tipo de lectura.
- *el referencial* está relacionado con una "estructura" construida: se tiene un conjunto coordinado de referentes. Los referenciales pueden darse fijos y de antemano en el proceso o permitir su construcción y reconstrucción.
- *la referencialización* –al decir de Figari (1993)– permite designar y poner de relieve al procedimiento mismo y significar que no puede tratarse ni de un objeto preexistente en relación con el cual los elementos de un dispositivo educativo tiene que situarse, ni de una norma o ley que fija finalidades establecidas. Este término lleva a una actividad que "consiste en emprender una búsqueda de referencias pertinentes [...], que puedan explicar y justificar la concepción y evaluación de un dispositivo educativo".

Es necesario señalar desde el comienzo quiénes son los destinatarios de la información y quiénes serán los encargados de concretar las acciones, y esto dependerá de que la evaluación esté orientada a la rendición de cuentas –en cuyo caso el eje es el control– o a mejorar el proyecto a partir de sus integrantes como protagonistas. Las propuestas de trabajo que se plantean a continuación están referidas exclusivamente a este último caso.

La información sobre el proyecto la brindan los actores directa e

indirectamente involucrados: alumnos, docentes, comunidad educativa, autoridades institucionales, etcétera; son aquellos que tienen algún interés en el mismo.

Se toma como punto de partida la consideración y la valoración de los acuerdos, las discrepancias, la identificación de los obstáculos y la elaboración de alternativas de acción.

Se propenderá a alentar el análisis, el debate y la indagación, lo que determinará la necesidad de un diseño abierto a posibles modificaciones. Los instrumentos de recolección de datos deben tener una estructura que permitan no sólo dar respuestas a categorías puntuales sino que reparen en lo singular, en lo imprevisible y en lo no pensado, lo que equivale a decir que proporcionen información tanto de los efectos buscados como de los no buscados, de estándares o de casos particulares.

Es éste el punto de mayor riqueza en la evaluación de un proyecto, ya que el docente, al salir de una actitud de constatación de los objetivos prefijados y tener la apertura para comprender los diferentes modos como los demás reciben el impacto de sus acciones, da lugar a conocer y comprender tanto los efectos buscados como los no buscados. Algunos de estos últimos pueden tener un gran valor educativo, otros pueden estar obstaculizando aprendizajes.

Así como la recolección, la organización y el primer análisis de los datos, son tareas casi exclusivas del docente, la interpretación puede ser compartida con otros participantes del proyecto (compañeros docentes, asesores pedagógicos, padres o alumnos avanzados según la edad). Son aquellos que a criterio del docente puedan generar sanos espacios de discusión.

Así, el recibir información sobre sus acciones en forma permanente le posibilita al docente mantener una vigilancia de sus prácticas en el aula y, desde allí, generar nuevos espacios de negociación que le permitan nuevas construcciones de aquéllas.

Propuesta de trabajo

Los interrogantes suelen ser un punto de partida para todo proceso evaluativo y en particular para la evaluación de un proyecto (en este caso el pedagógico-didáctico del docente), que básicamente podría dirigir la mirada sobre cuestiones generales. No podemos

desconocer que detrás de una pregunta existe una afirmación hipotética que tiene que poder "leerse" desde algún lugar.

¿Qué función cumple el proyecto pedagógico-didáctico que lleva adelante el docente en el marco de la formación del estudiante? ¿Cómo se articula su propuesta con la función social de la escuela? ¿Qué lugar ocupa la escuela en la formación del ciudadano y, desde allí, cómo se articula su propuesta? ¿Cuáles son las demandas de formación que recibe el docente y cuáles de ellas responden a genuinos intereses comunitarios o a intereses personales o corporativos de sectores?

Pero, aunque los interrogantes son esenciales para una tarea evaluativa, donde se interpelen los fundamentos y las estrategias de indagación de la realidad, la evaluación no se hace simplemente dando una respuesta a aquéllos, sino que requiere la construcción de procedimientos de referencialización que den cuenta de estructuras organizadoras basadas en sistemas de referencias.

Se presenta un cuadro, sólo a modo de ejemplo, que permita aproximarse a una referencialización:

Dimensiones	*Categorías*	*Indicadores*
Propuesta pedagógico-didáctica	Relación teoría-práctica	* La fundamentación de las acciones llevadas a cabo en el aula.
	Tratamiento de contenidos	* La actualización de conocimientos en el desarrollo del campo científico.
		* Las posibilidades de generar espacios de discusión.
Organización académico-administrativa	Organización del tiempo y del espacio	* Distribución horaria.
		* Recursos requeridos.

Se ha tenido especial cuidado de que el cuadro no haga perder la visión de totalidad, ya que el recorte fragmentado atomiza, dificultando una tarea de reestructuración permanente, propia de la referencialización.

La metodología y los instrumentos de recolección de datos pueden ser:

- observación de clases;
- discusión con un par de la misma área o grado;
- discusión en reuniones de personal;
- entrevistas;
- encuestas;
- carpetas de los alumnos;
- reuniones de padres;
- las propias carpetas didácticas o planificaciones.

Los informantes, actores directa e indirectamente involucrados, son:

- director;
- alumnos del mismo curso;
- alumnos que ya cursaron;
- alumnos que aún no han cursado;
- padres;
- supervisor;
- pares docentes;
- personal administrativo;
- psicopedagogo o grupo técnico relacionado con el curso;
- el propio docente a través de la autoevaluación.

La recolección de la información puede hacerse sobre la base de entrevistas, encuestas anónimas, discusiones grupales, informantes clave, observaciones, grabación de clases, trabajos de los alumnos, así como también las clásicas evaluaciones de los aprendizajes utilizadas para acreditar.

En el análisis de los datos se tendrá un particular cuidado de referencializarlos inicialmente a partir de las características del actor-informante que los proporcione; así, por ejemplo, el psicopedagogo tendrá como sistema de referencias –que es donde apoya su

mirada selectiva– el tipo de formación que haya tenido; y el director, al igual que el supervisor, tendrá un enfoque de gestión. Los actores directa e indirectamente involucrados tienen diferentes niveles de significación de acuerdo con los referenciales construidos previamente.

Esta tarea, que seguramente genera confrontaciones y discusiones –allí radica gran parte del potencial de los aportes–, se puede hacer habitual; el docente se integra con mayor facilidad a la institución, ya que tanto las críticas de sus pares como los señalamientos de los niveles directivos pueden resignificarse como la mirada que tiene otro en relación con su práctica o, dicho de otra manera, cómo son reinterpretadas las acciones del docente en su comunidad, quedando también claro que las estructuras particulares de cada institución configuran y limitan su acción.

En todo este proceso no se puede desconocer que tomar conciencia del contexto institucional requiere de procesos cognitivos que deben ser desarrollados. Mantener un equilibrio entre la reflexión del sujeto docente y las miradas de otros integrantes que constituyen el proyecto pedagógico-didáctico podría permitir un accionar informado, sujeto a una permanente reconstrucción.

La autoevaluación

La autoevaluación en el ámbito de la educación, hasta no hace mucho tiempo, se concebía como un valor en sí misma, generalmente incuestionable. Pero luego de que fuera utilizada por los organismos internacionales como proceso previo a la evaluación externa se la percibe hoy con cierta desconfianza. Sin embargo, para su análisis sería importante distinguir desde qué proyecto se plantea, para poder asignarle el lugar y el valor que le corresponden y no caer en espacios vacíos, cargados de palabras ideologizadas que entorpecen aún más las prácticas escolares.

Es decir que, según desde dónde surja el requerimiento de autoevaluación, puede estar directamente ligada a los sistemas de control y/o cruzada por una perspectiva de control. Si no se explicita este lugar la autoevaluación funciona "como si", ya que resultaría torpe pensar que un docente o un miembro de la comunidad educativa

pueden mostrar aquellos aspectos de su accionar que no favorezcan su imagen.

Desde este lugar se puede hablar –y desde allí se debiera interpretar– una autoevaluación interferida, cruzada por un sentido –favorable o desfavorable– de autoridad como legitimadora o deslegitimadora de saberes y acciones.

Sin embargo, un sistema de autoevaluación que podría llamarse interferida –desde un sentido clásico del término– permite ir construyendo acuerdos básicos, a partir de la explicitación de intereses y valores, que muchas veces se dan por supuestos. Así al explicitar las "reglas de juego" pueden trasparentarse las relaciones y funciones de los actores que intervienen en el proyecto.

En el proyecto pedagógico que el docente proponga como propio, donde él decida las acciones por seguir, una mirada hacia el interior cobra un sentido de naturaleza diferente, lo lleva a volver a recorrer un camino en un intento de comprensión histórica. También una mirada de reconceptualización de su tarea, y reconocer los entrecruzamientos epistemológicos, éticos y políticos como parte constitutiva de su propia práctica; es el tomar distancia de la práctica para trabajar sobre ella, objetivar los compromisos con la lectura de la propia práctica, así como también observar las propias posibilidades y limitaciones y desde allí poder ver lo que en la acción es imperceptible.

La autoevaluación así concebida debe tener –como mínimo– dos condiciones básicas: una relacionada con la capacidad de objetivar las acciones realizadas, y la otra ligada a la responsabilidad y al compromiso.

En el primer caso es necesario poner distancia como si las acciones que uno mismo realiza fueran hechas por otro y, mejor aún, compararlas con prácticas similares en otros ámbitos culturales. En el segundo caso esta actitud de "extrañamiento" no debe ser tomada como anécdota sino como desafío para mejorar las prácticas.

En síntesis, la autoevaluación lleva implícita un proceso reflexivo, que tiene diferentes formas de ser analizado, de acuerdo con la naturaleza de las metas que orientan la acción, es decir, de acuerdo con el proyecto en el que se inscribe.

Acerca del docente

El docente es un sujeto social, y desde allí construye su profesionalidad. Tomando este punto de partida se concibe al maestro como un profesional que, además de apropiarse del conocimiento científico y tecnológico, puede ubicarse frente a su propia práctica reflexionando acerca de los supuestos que dan sentido a su tarea en relación a para qué, qué y cómo enseña, cómo aprenden sus alumnos y cómo promueve la relación con el conocimiento.

La profesionalización implica –además del dominio de la disciplina– la toma de conciencia de las propias contradicciones que enmarcan su práctica cotidiana. Al decir "contradicciones" se hace referencia a la inconsistencia, al quiebre, al espacio de fractura entre lo pensado, lo dicho y lo actuado. Avanzar en el conocimiento de dichas contradicciones es detenerse a reflexionar[16] sobre qué se hace, cómo se hace y para qué se hace.

Una de las primeras miradas, y quizá la más compleja, es el intento de desentrañar el marco ideológico-axiológico que orienta la acción del docente. Los modos de pensar, las formas de comprender la vida, los valores implícitos son diferentes en cada docente. Cada uno responde a una cosmovisión en la que inciden su pertenencia social y su peculiar historia de vida.

Para pensar en un proceso de autoevaluación del docente es importante señalar la función del metaanálisis:

> a la manera de una transparencia o un cristal que permite una nueva mirada sobre la clase. La superposición de las transparencias nos remite a una narrativa que deja de hablar de la clase vivida para construir un

16. Según Elliot (1993) es posible distinguir dos tipos muy diferentes de desarrollo reflexivo de la práctica docente.

1) El profesor emprende una investigación sobre un problema práctico, cambiando sobre esta base algún aspecto de su práctica docente. El desarrollo de la comprensión precede a la decisión de cambiar las estrategias docentes. En otras palabras, la reflexión inicia la acción.

2) El profesor modifica algún aspecto de su práctica docente como respuesta a algún problema práctico, revisando después su eficacia para resolverlo. Mediante la evaluación, la comprensión inicial del profesor sobre el problema se modifica y cambia. Por lo tanto, la decisión de adoptar una estrategia de cambio precede al desarrollo de la comprensión. La acción inicia la reflexión.

> nuevo relato que cobra vida independientemente de ella. Cada una de las construcciones permite superar los niveles interpretativos del mismo contenido tratado y poner énfasis en los aspectos substanciales [...] se favorecen reflexiones sobre cada una de las decisiones que se adoptan para la enseñanza (Litwin, 1993).

Su actitud frente a los obstáculos puede hacerlo recurrir a un pensamiento normativo, racionalizador o elucidante. En el primer caso, son las normas vigentes las que justifican la situación; en el segundo, la causa da cuenta y agota el problema; en cambio, en el último caso –pensamiento elucidante– se trata de dilucidar los conflictos transformándolos de dilemas en problemas, de manera tal que permita buscar nuevas formas de resolverlos.

No le es simple al docente poder objetivar su práctica para tomar una actitud problematizadora frente a la realidad, ya que existen mecanismos justificatorios que atraviesan la cotidianidad del maestro y que pueden tapar los condicionantes de su práctica. En este sentido resulta valioso resaltar el planteo realizado por Elena Achili (1991), quien a partir de una experiencia realizada en talleres de educadores señala algunas características del pensamiento del docente que ella misma refiere a explicaciones desde una construcción social de la imagen del docente.

Afirma esta investigadora que se parte de la idea de que tanto los saberes como las representaciones de los docentes son heterogéneos y que poder explicitarlos puede colaborar a señalar obstáculos en la construcción del conocimiento. Es posible encontrar un "...cierto pragmatismo y espontaneísmo, una inclinación hacia lo particular, un propender a la evaluación/valoración y una desvalorización de su propio saber".

En relación al pragmatismo hay una ligazón a lo empírico, a lo utilitario que le resta libertad para realizar articulaciones y conexiones más generales. Se produce una indiferenciación entre la reflexión sobre la práctica y la práctica misma.

Existe un particularismo de rápida generalización en la medida en que los saberes generados en una experiencia individual, como respuesta a situaciones particulares, se cargan de la normatividad vigente y se generalizan.

Una mirada reflexiva y crítica de sus saberes puede sacarlo de la

búsqueda permanente de legitimación externa de modo de entender su desvalorización por una construcción social, y en lugar de buscar una convalidación por valoración de saberes, poder encontrar en su acción la valoración o desvalorización que él hace en el aula. Él busca que le enseñen y le resulta muy difícil aprender de sus pares, más bien compite con ellos.

La actitud valorativa perturba la descripción de situaciones e impide el análisis, y a veces cristaliza las situaciones dilemáticas. Esto tiene que ver con su formación y con su relación institucional, donde es observado y evaluado constantemente.

¿Cómo puede evaluar el docente su propia práctica?

En primer lugar, debe tener como punto de partida una propuesta de trabajo en el aula de carácter experimental, ya que para poder reflexionar tiene que existir una instancia previa de hipotetización. A partir de una adhesión provisional de la experiencia se provoca una crítica a las reglas prescriptas, en una actitud permanente que incluya la aceptación o el rechazo de los resultados, en donde se consolidan las capacidades reflexivas, donde se promueve un pensamiento divergente.

La aceptación de las diferencias entre las aspiraciones y las prácticas reales que se dan, en una actitud objetiva, es el punto de partida de la tarea. "Los docentes deben tener capacidad para tolerar las pérdidas de autoestima, que suelen acompañar a la autoinspección" (Elliot, 1993).

Esta pérdida de autoestima sólo puede ser entendida si su práctica se valoriza desde su capacidad metacognitiva, de la ponderación de los análisis que surgen de la reflexión con valoración de ese proceso. Es muy difícil aceptar desde una mirada tradicional, donde el valor de la práctica de los docentes está totalmente ligado a los resultados de los aprendizajes de los alumnos, la falta de éxito en este sentido.

Para una tarea de autoevaluación parece interesante el planteo que hace Telma Barreiro (1988), quien realiza un esquema del accionar del docente en el que marca niveles; dos de ellos pueden aportar elementos para el análisis:

1. Un nivel manifiesto o explícito, donde se hallan las conductas observables que un profesor o maestro tiene en su actividad.

2. Un nivel subyacente o implícito donde se puede encontrar:

a) Su bagaje técnico o profesional (conocimientos y habilidades técnicas específicas de su especialidad).

b) Supuestos, creencias y valores de carácter filosófico, pedagógico o sociológico, que constituyen una visión del mundo.

Un muy buen punto de partida para la autoevaluación es encontrar las consistencias o contradicciones que existen entre niveles, partiendo de la premisa de que las contradicciones forman parte de un proceso de construcción cuando se puede dar cuenta de ellas y encontrar modos que generen nuevas propuestas o señalen nuevas contradicciones.

Esta mirada interna se vuelve estéril si no se logra encontrar un "otro" con quien se pueda monitorear los procesos internos. Para que esto suceda es necesario que se den determinadas condiciones:

1. *Confidencialidad*: sólo una sincera relación lleva a confrontar en un mismo nivel los diferentes modos de pensamiento. Los mecanismos de defensa como protección obstaculizan y desnaturalizan los procesos reflexivos.

2. *Confianza*: desde este lugar hay una gran tendencia a seleccionar los datos positivos. Trabajar con un igual permite un verdadero intercambio.

3. *Reciprocidad*: "...sólo con tiempo, confianza y reciprocidad puede medirse y aceptarse de buena fe el nivel apropiado de intercambio" (Mortimore, 1986).

Con este punto de partida el profesor puede desarrollar su conocimiento práctico sobre las situaciones educativas en el aula, superando tanto las deficiencias del conocimiento profesional tácito acumulado a lo largo de la experiencia como el conocimiento académico descontextualizado, incapaz de salvar el abismo entre la teoría y la práctica.[17]

17. Kemmis, en el prólogo del libro de Carr (1990) *Hacia una ciencia crítica de la educación*, marca con claridad la relación teoría-práctica: "...es siendo teorizadas

La presencia de un "otro" se plantea también como necesaria para no desarrollar teorías a partir únicamente de la reflexión sobre la experiencia, ya que la reflexión aislada corre el riesgo tanto de generar y reproducir autocomprensiones deformadas de la realidad como de destinar grandes esfuerzos en descubrir cuestiones que ya fueron trabajadas por otros. Al respecto, Elliot (1993) señala que dejar de lado las reflexiones de los demás podría llevar al docente al gran descubrimiento de inventar la rueda.

Desde este lugar se puede afirmar que el conocimiento científico y cultural acumulado en la historia de la humanidad y de la profesión en particular nunca podrá sustituir la reflexión de los profesores, pero sí es un elemento imprescindible para apoyar esta reflexión.

El valor de tomar una nueva mirada ayuda a comprender a través del "otro", a clarificar nuestros marcos ideológicos que obturan nuestra mirada de la realidad y a encontrar nuevos "modos" de leerla.

No se pueden valorar los procesos reflexivos en relación con una práctica si no se tiene presente que junto con la actitud reflexiva debiera –equilibradamente– estar presente una actitud de búsqueda de modos alternativos de actuar, búsqueda que articule la interrogación con las respuestas provisionales, con la precaución de no sobrevalorar lo nuevo en detrimento de lo clásicamente valioso.

Para organizar un sistema de referencialización puede ser un ejemplo el siguiente cuadro:

---

que las prácticas tienen sentido (como prácticas de cierta clase), y es siendo practicadas que las teorías tienen importancia histórica social y material. La teoría no consiste solamente en palabras y la práctica no es solo mudo comportamiento; la teoría y la práctica son mutuamente aspectos constitutivos uno de otro. Bajo este punto de vista, no puede haber ningún 'abismo' entre la teoría y la práctica, solamente mayores o menores grados de desarmonía, elisión e ilusión en la relación recíproca. Solamente podemos identificar estas desarmonías, elisiones e ilusiones examinando cómo se relacionan entre sí nuestras teorías y prácticas...".

Dimensiones	*Categorías*	*Indicadores*
Conocimiento	Actitud frente al conocimiento	En relación con – el tipo de planificación – la revisión de lo planificado: * fundamentar lo actuado * proponer varias alternativas de acción
Relación docente-alumno	Actitud frente al alumno	–Comprensión de su lógica particular –Interpretación de los requerimientos –Alerta sobre su propia actitud de censura
	Actitud frente al grupo	–Respeto por acuerdos contractuados –Alerta ante las posibilidades de negociación de significados que brinda el grupo

Condiciones institucionales para la autoevaluación del docente

El punto de partida para comprender el tipo o las características que asumirá la autoevaluación del docente, es analizar la persona u organismo que propone la autoevaluación. Si ésta surge de una demanda gubernamental o desde la misma institución en un espacio burocratizado, tal como se señaló anteriormente, se tratará de una autoevaluación interferida, cuyo valor puede darse por la generación de espacios de negociación. La autoevaluación debe comprenderse en el marco del proyecto en el que se inscribe, directamente referenciada a las metas y a los supuestos ideológicos y axiológicos que lo enmarcan, con lo que implica pasar la información del orden privado al público.

Para que haya una autocrítica por parte del docente, es necesario que dentro de la institución estén dadas las condiciones para ella. El

punto indiscutible es que todos los sectores que integran la institución, incluido las autoridades, realicen su propia autoevaluación, cuidando siempre, tanto desde los docentes como de los mismos directivos, que no se transforme en un espacio de "pase de boletas" que sólo conduce a mecanismos defensivos.

En este sentido no se puede desconocer la representación que tienen los docentes de la autoridad escolar, ya que su función histórica de control y sanción ha generado en ellos mecanismos de defensa que los llevan a mostrar sus buenas prácticas y a esconder aquellas que no tuvieron el "éxito" esperado.

Uno de los mayores riesgos es que se instale una falsa democratización de las decisiones y el accionar se vuelva intransitivo y vacío. Desde este lugar sólo enmascara contradicciones. Si por el contrario se discuten distintos planteamientos con sucesivas autoevaluaciones, explicitando sus dificultades, se puede ayudar al docente a tener una mayor seguridad en lo que hace.

Es necesario ir desconstruyendo la historia del control y de sanción que caracterizó a la escuela, para generar espacios sanos de interacción docente.

Telma Barreiro (1988) analiza las características de lo que ella define como grupo sano, y si bien su propuesta hace referencia a pequeños grupos cuyas relaciones personales están ligadas por lazos afectivos especiales, algunas de las condiciones señaladas como necesarias se pueden dar dentro del ámbito de una escuela o, al menos, ser motivo de reflexión.

Según esta autora, el grupo promueve el crecimiento de sus miembros, teniendo como pilar importante la seguridad básica.

- No hay autoridad "autoritaria", sino que todos los integrantes del grupo aceptan e incorporan críticas, y fundamentan sus pedidos.
- Hay respeto por la condición humana de sus integrantes, el grupo valora los conocimientos y las aptitudes de aquéllos, por lo que ningún miembro se siente marginado.
- Existe confianza entre sus miembros.
- Se acepta que cada miembro tiene conflictos, necesidades que parten de su realidad e historia personal.
- No se teme sacar a luz los conflictos, se discute con sinceridad y,

lo que resulta más importante, con una actitud de apertura, donde desaparece el que gana y el que pierde, se trata de expresar las dudas cuando se presentan, sin esperar que socaven las relaciones.
- Se trata de evitar los códigos dobles, aceptándose los puntos de vista individuales, a la vez que se intenta lograr coherencia en los valores.
- Interesan las cuestiones de fondo, se descartan las cuestiones formales.
- Sus miembros no se preocupan por cuidar una imagen; desestiman que sus acciones se vean como defectos, sino más bien como consecuencia de su modo de pensar.
- El grupo busca trascender hacia afuera, no es autosatisfactorio, se integra a la comunidad sin generar competencias planteando acciones complementarias en función de un bien común.

Todas estas características se señalan como importantes para una interacción entre los integrantes de una escuela, desde donde se percibe la posibilidad de ir construyendo un "clima institucional" que permita la autoevaluación.

Cabría señalar que es imprescindible que sus miembros puedan diferenciar los espacios públicos de los privados, ya que los diferentes escenarios transforman las acciones, pudiendo considerarse "traiciones" los relatos que circulan sin tener el sentido que se les quiso dar.

En los momentos de cambios curriculares, no puede olvidarse que son esencialmente procesos sociales y por tanto tan complejos como lo son todos los procesos de este tipo.

> En su interior se encuentran polémicas, estructuras y superestructuras, contradicciones y conflictos. Los procesos de determinación social son aquellos en los cuales a través de luchas, negociaciones o imposiciones en un momento de transformación o génesis se producen rasgos [...] que van a configurar una estructura social relativamente estable y que tienden a definir límites y posibilidades [...] (De Alba, 1995).

Si las primeras experiencias no resguardan estos señalamientos es muy probable que los procesos de autoevaluación queden truncados.

Autoevaluación del alumno

Dos dimensiones parecen presentarse en el análisis de la autoevaluación del alumno, una referida a su relación con los pares y con el maestro, y otra, a la búsqueda de indicadores que le permitan conocer sus propios procesos cognitivos, o los modos o maneras de resolver una situación y los modos o maneras de apropiación del conocimiento. En el primer caso daría cuenta de su lugar frente al grupo de pares, en el reconocimiento de su lugar para favorecer el "clima" de trabajo en el aula. En el segundo caso estaría ligado a un autoconocimiento, a lo que podríamos llamar una metacognición.

Al hacer referencia al "clima" dentro del aula, son varios los aspectos que pueden o deben ser considerados, ya que pareciera, desde la construcción histórica de lo que "debe" ser un aula, que el solo hecho de no tener desorden o hacer silencio fuera suficiente. Desde este lugar podría analizarse la función del contrato didáctico[18] –donde interjuegan lo aceptable y lo valioso– para la delimitación y las transacciones de las relaciones y los requerimientos entre docentes y alumnos en función de la apropiación de conocimentos por parte de estos últimos.

Es en el espacio del aula donde se producen intercambios de significados, y es en la intención entre pares donde se da gran parte de la construcción de conocimientos, sobre todo si se tiende a formar en el alumno un pensamiento divergente, en el cual el docente no distribuye verdades sino que genera espacios de discusión sobre propuestas hipotéticas. Es a partir del compromiso y el aporte a la construcción colectiva que puede tomarse la autoevaluación del alumno.

Un muy buen punto de partida para la autoevaluación es la tarea de reconocimiento de las prácticas de sus pares que han colaborado en los procesos de construcción compartida; de esta manera se sale de la idea inicial de señalamiento de las dificultades que pueden

18. Acuerdo entre docente y alumnos directamente relacionado con un contenido disciplinar específico, a diferencia del contrato pedagógico de carácter más general, que regula las interacciones de docentes y alumnos que no dependen de un contenido, y del contrato escolar que gobierna la institución escuela.

generar irritaciones y mecanismos defensivos, dando lugar a efectos contrarios a los que se buscan. Desde aquí se favorece la construcción de la autoimagen dentro del grupo, así como también permite una relativización de las dificultades cuando éstas aparecen. Desde este lugar la ubicación dentro del grupo le permite a cada alumno reconocer los aprendizajes que le posibilitó el intercambio grupal, viendo desde su lugar qué aporta y qué recibe del grupo.

Así como la autoevaluación dentro de un grupo no puede perder de vista la importancia de la negociación de significados entre pares, la autoevaluación desde una mirada individual debiera implicar un reconocimiento de las habilidades cognitivas.

Ambos casos –la autoevaluación grupal y la metaevaluación– tienen como eje la construcción del conocimiento por parte del alumno; lo que marca la mirada en común es la actitud de provisionalidad frente al conocimiento.

En este sentido el punto de partida parece ser el reconocimiento de lo que el sujeto está haciendo, el reconocimiento de las estrategias intelectuales[19] que se ponen en juego para abordar una tarea desconocida.

Si se analizan las acciones que los niños realizan en el aula, hasta se puede visualizar cuando borran en el transcurso de la ejecución de una actividad, lo que está mostrando que paralelamente a la ejecución hay una actividad de control, que regula y ajusta en función de la representación que el niño tenga de lo que debe lograr.

Este punto de la representación de los logros parece ser crucial en el diseño de los planes de acción que el alumno pone en juego en la ejecución de la tarea, de allí la importancia de tener claros los niveles de desempeño que debe lograr en tiempos largos o cortos, según la tarea que emprende.

Nunziati (1990) trabaja una propuesta pedagógica centrada en la autoevaluación poniendo especial énfasis en la importancia de la autorregulación del alumno de sus propios procesos de aprendizaje, basándose en la posibilidad de que él pueda tener una repre-

19. En el sentido que Nisbet y Schucksmith (1987) dan a "estrategias": "Secuencias integradas de procedimientos que se eligen con el propósito de facilitar la apropiación de conocimiento, lo que nos permite utilizar información de manera precisa y en el momento oportuno".

sentación de antemano de éstos, que en definitiva son los criterios de valoración que el docente tiene de su tarea. Este autor considera que "...la autoevaluación es una 'habilidad' a construir. Pasar de una fase evaluativa espontánea a un sistema de pilotaje de experto es el resultado de un aprendizaje".

En este proceso se pueden encontrar dos fases; la primera es de apropiación de criterios, ya que es una fase de organización del funcionamiento de la autoevaluación. Se aproxima a los criterios definidos, puntuales y detallados, y a las operaciones que componen el saber-hacer.

La segunda se refiere a las actividades de autoevaluación propiamente dichas y que implican la elección del momento cuando son exigidas, la forma como se realizarán según lo que se piensa hacer. Esto puede obligar a una reformulación de criterios y a una comparación real del trabajo realizado.

La importancia de la autoevaluación consiste en la verbalización interna que ella impone, que muchas veces se omite, así como también la reformulación de criterios que obturan la construcción y organización de los conocimientos.

S. Tishman, D. Perkins, y E. Jay, (1997) otorgan un lugar importante a los procesos metacognitivos en la formación del pensamiento. Los autores señalan seis dimensiones del "buen pensamiento"[20] donde se resalta:

- el lenguaje de pensamiento, relacionada con los términos y conceptos que se usan en el aula para referirse al pensamiento;
- las predisposiciones al pensamiento, lo que marca actitudes, valores y hábitos;
- el monitoreo mental, referida a la reflexión de los alumnos sobre sus propios procesos de pensamiento para que tomen el control de una manera creativa y eficiente;
- el espíritu estratégico, que marca una actitud especial para construir y usar estrategias de pensamiento como respuesta a desafíos intelectuales;

20. Es muy interesante el planteo de los autores en relación con la importancia de generar una "cultura de pensamiento" dentro del aula, donde las dimensiones que se desarrollaron operan como sugerencias didácticas para un trabajo en el aula.

- el conocimiento de orden superior que intenta superar el conocimiento fáctico de una materia, y
- la transferencia, referida a la aplicación de conocimientos y estrategias de un contenido en otro.

En general los alumnos no son conscientes de lo que hacen ni de sus estilos de aprendizaje, lo que podría lograrse a través de procesos metacognitivos que brindan la capacidad de conocer el propio conocimiento, de pensar y reflexionar sobre cómo reaccionamos o hemos reaccionado frente a una tarea o problema.

Flavell[21] describe con claridad la metacognición como

> el conocimiento de uno mismo concerniente a los propios procesos y productos cognitivos o a todo lo relaciondo con ello, por ejemplo las propiedades de información o datos relevantes para el aprendizaje. Así practico la metacognición (metamemoria, metaaprendizaje, metaatención, metalenguaje, etcétera) cuando caigo en la cuenta de que tengo más dificultad en aprender A que B, cuando comprendo que debo verificar por segunda vez C antes de aceptarlo como un hecho; cuando se me ocurre que haría bien en examinar todas y cada una de las alternativas en una elección múltiple antes de decidir cuál es la mejor; cuando advierto que debo tomar nota de D porque puedo olvidarlo [...] La metacognición indica, entre otras cosas, el examen activo y consiguiente regulación y organización de estos procesos en relación con los objetos cognitivos sobre los que versan, por lo general al servicio de algún fin u objetivo concreto.

La toma de conciencia de su modo de aprender y de la complejidad del mismo son fundamentales para poder controlar su aprendizaje y, desde allí, planificar y organizar sus propias actividades de aprendizaje, tratando de gestar una disposición habitual que pueda ser intrínsecamente provechosa, donde se articulen armónicamente la reflexión, la interpelación y la imaginación.

Se presenta un ejemplo que ayude a referencializar la auto-evaluación en relación a las acciones de estudio, entendiendo que para esta tarea existe por parte del alumno una planificación de acciones y una definición de sus objetivos.

21. Citado por Nisbet, J. y Schucksmith, J. (1987).

Dimensión	*Categoría*	*Indicadores*
Planificación de las acciones	Definición de objetivos	– Claridad en la delimitación – Secuenciación temporal – Actitud de duda – Apertura a una nueva organización
	Estrategias utilizadas	– Lectura general o parcial del material – Relectura de obras complejas – Fijación de palabras que representan conceptos clave – Identificación del libro o la persona que puede proporcionar información aclaratoria

Un ejemplo de referencialización en el área de la comprensión relacionada con la evaluación de logros:

Dimensión	*Categoría*	*Indicadores*
Comprensión	Lectura	– Reordenar explicaciones – Encontrar ejes – Precisar vocabulario
	Explicación teórica	– Encontrar un sentido – Enmarcar posicionamiento teórico – Resignificación de conocimientos previos – Presentar nuevos interrogantes

Los procesos de construcción son diferentes en las distintas áreas de conocimiento, por lo que es muy importante establecer con los alumnos cuáles son los indicadores que dan cuenta del saber y del saber hacer en cada campo disciplinar. Esto permite revisar los criterios y superar en parte los referentes impuestos.

En todo proceso de autoevaluación, las acciones realizadas a

partir de la valoración de los datos son constitutivas del mismo proceso, de lo contrario éste perdería su sentido.

Acciones del docente que pueden favorecer los procesos de autoevaluación del alumno

Es importante ayudar al alumno a generar una actitud de disconformidad cuando no comprende una explicación o un texto no le resulta claro. Esta ayuda no está referida a la sanción sino más bien ligada a señalar la importancia de las nuevas relaciones que proporciona un conocimiento trabajado y asumir una actitud de búsqueda de los factores que interfieren su comprensión. Esta actitud de "batallar" con el conocimiento requiere que el docente demande de sus alumnos una fundamentación permanente de sus afirmaciones. Esto coloca al estudiante como sujeto que interpela permanentemente al texto.

El método socrático es un modo de actuación docente que coloca al alumno frente al conocimiento. Esto se puede efectivizar a través de una autointerrogación permanente y en el reconocimiento de los saberes previos en relación con el contenido del material por trabajar.

Una etapa de intercambio entre compañeros y con los mismos docentes, de procesos internos de autoevaluación, es de gran valor. Una actitud de cooperación, no de competencia, puede aportar categorías e indicadores que ayuden al alumno a repensar su aprendizaje, a partir de los modos de aprender de los demás. El abanico de posibilidades que se presenta en los modelos permite a los alumnos la apropiación de aquellos que pueden estar de acuerdo con sus características personales.

Puede considerarse también una buena estrategia que el maestro genere espacios grupales donde los alumnos puedan verbalizar las estrategias puestas en juego ante un problema o para efectivizar una actividad. En esta tarea resulta vital que el mismo docente explique las acciones que puso en juego.

Ian Selmes (1988) propone acciones concretas para ayudar al alumno en el conocimiento de su propio aprendizaje. Su propuesta, denominada PER, implica propósito, estrategias y revisión:

el propósito hace referencia a la percepción de los alumnos en cuanto al objetivo y a la importancia de una tarea específica; la estrategia se relaciona con el modo en que se organiza y realiza la tarea; la revisión tiene que ver con la identificación del resultado al haber utilizado la estrategia y su comparación con el propósito para decidir si la tarea ha sido realizada con éxito o no...

Según este autor, la clara definición del propósito de la lectura, que está marcado por la intencionalidad, determina el tipo de conocimiento: superficial o profundo. En el primer caso los propósitos están referidos principalmente al cumplimiento de los requisitos de la tarea; en el segundo, el propósito es comprender el significado de lo que se lee.

Entwistle (1991) también plantea una relación entre el propósito de la lectura, a lo que agrega la motivación y los procesos:

Motivación	*Intención*	*Enfoque*	*Procesos*
Intrínseca	Comprender	Profundo	Relacionar con la experiencia y los conocimientos previos
Miedo al fracaso	Cumplir los requisitos de la tarea	Superficial	Memorizar temas sueltos de información
Necesidad de rendimiento	Obtener notas lo más altas posible	Estratégico	Asignar tiempo, esfuerzo y enfoques según la "rentabilidad"

La autoevaluación y el proyecto pedagógico-didáctico

Encontrar en el ritmo que vive hoy la escuela un espacio de reflexión resuena más a utopía que a posibilidad, pero es conveniente que si el docente visualiza propuestas que pueden mejorar su práctica las lleve adelante en forma provisional, tan sólo con algún grupo de alumnos y en tiempos acotados. Esta actitud de prueba puede transformarse en una verdadera hipótesis de trabajo, al

permitir permanentes revisiones y cuestionamientos constructivos que pueden dar lugar a economía de esfuerzo y tiempo para ser trabajados con otros grupos.

No se puede desconocer que los modelos docentes han dado respuestas a requerimientos de la práctica que hoy no son suficientes; es por eso que una revisión nos obliga a mejorar la capacidad de "escucha" tanto en el "reconocimiento" de nuestras prácticas como en el comprender al "otro" desde sus marcos interpretativos, para desde allí generar cambios.

Tomar la autoevaluación como eje de análisis y su articulación con el proyecto pedagógico-didáctico marca un posicionamiento que concibe tanto al docente como al alumno como protagonistas de sus propias decisiones. Esto implica aceptar, en cierta medida, que dentro de las limitaciones que el contexto marca existen espacios que permiten acrecentar la autonomía. Ofrecer elementos concretos que permitan una mirada totalizadora de la realidad no significa quedarse en lo instrumental, sino más bien posiciona al docente (con instrumentos) frente a los proyectos macro que contextualizan el quehacer de la escuela.

Bibliografía

Achilli, Elena Libia: *La práctica docente: una interpretación desde los saberes del maestro*, Cuadernos de Formación Docente, Universidad Nacional de Rosario, nº 1, 1991.

Allal, Linda: "Estrategias de evaluación formativa: concepciones psicopedagógicas y modalidades de aplicación", *Infancia y aprendizaje*, nº 11, Madrid, 1980.

Álvarez Méndez, Juan Manuel: "El alumnado. La evaluación como actividad crítica del aprendizaje", *Cuadernos de Pedagogía*, nº 219, 1993.

Ardoino, Jacques y Berger, Guy: *Evaluation en actes*, mimeografiado, 1988.

Ardoino, Jacques: "Las posturas (o imposturas) respectivas del investigador, del experto y del consultor", en Ducoing, P. y Landesmann, M. (compiladoras y editoras), *Las nuevas formas de*

investigación en educación, Ambassade de France au Mexique-Universidad Autónoma de Hidalgo, México DF, 1993.

Barreiro, Telma: "Los grupos de reflexión, encuentro y crecimiento (GREC): Una propuesta para el perfeccionamiento docente", *Revista Argentina de Educación*, año VI, nº 11, diciembre de 1988.

———: *Trabajos en grupo*, Buenos Aires, Kapelusz, 1992.

Bernstein, Basil: *Poder, educación y conciencia. Sociología de la transmisión cultural*, Barcelona, El Roure Editorial S.A., 1990.

Bertoni, Alicia; Poggi, Margarita, y Teobaldo, Marta: *Evaluación. Nuevos significados para una práctica compleja*, Buenos Aires, Kapelusz, 1995.

Camilloni, Alicia W. de: "La evaluación como una operación de construcción y de comunicación de un juicio de valor", *Novedades Educativas*, nº 57, 1995.

Carr, Wilfred: *Hacia un ciencia crítica de la educación*, Barcelona, Editorial Laertes, 1990.

Casanova, María Antonia: *Manual de evaluación educativa*, Madrid, La Muralla, 1995.

De Alba, Alicia; Díaz Barriga, Angel, y Viesca, Marta: "Evaluación: Análisis de una noción", *Revista Mexicana de Sociología*, año XLVI, vol. XLVI, nº 1, 1988.

De Alba, Alicia: *Currículum: Crisis, mito y perspectivas*, Buenos Aires, Miño y Dávila, 1995.

Dewey, John: *Cómo pensamos*, Barcelona, Paidós, 1989.

Díaz Barriga, Ángel: *Didáctica y currículum*, México, Editorial Nuevomar, 1985.

———: *Docente y programa. Lo institucional y lo didáctico*, Buenos Aires, Aique, 1994.

Edelstein, Gloria y Coria, Adela: *Imágenes e imaginación. Iniciación a la docencia*, Buenos Aires, Kapelusz, 1995.

Elliot, John: *La investigación acción en educación*, Madrid, Morata, 1990.

———: *El cambio educativo desde la investigación acción*, Madrid, Morata, 1993.

Entwistle, Noel: *La comprensión del aprendizaje en el aula*, Barcelona, Paidós, 1991.

Fígari, Gerard: "Quel système de références pour évaluer une mesostructure éducative?", en Masse, Denis, *L'evaluation institutionnelle en milieu scolaire*, París, 1993.

Gardner, Howard: *Inteligencias múltiples. La teoría en la práctica*, Barcelona, Paidós, 1995.

Gimeno Sacristán, José y Pérez Gómez, Ángel: *Comprender y transformar la enseñanza*, Madrid, Morata, 1992.

———: *La enseñanza, su teoría y su práctica*, Madrid, AKAL, 1985.

House, Ernest: *Evaluación, ética y poder*, Madrid, Morata, 1994.

———: "Tendencias en evaluación", *Revista de Educación*, nº 299, Madrid, 1992.

Jackson, Phillips: *La vida en las aulas*, Madrid, Morata, 1992.

Jorba, Jaume y Casellas, Ester: *La regulació y l'autoregulació dels aprenentatges*, Universidad Autónoma de Barcelona: Institut de Ciencies de l'Educació, 1996.

Lafourcade, Pedro: *La autoevaluación institucional*, Buenos Aires, Kapelusz, 1992.

Landsheere, Gilbert: *Evaluación continua y exámenes*, Buenos Aires, El Ateneo, 1978.

Litwin, Edith: "Algunas reflexiones en torno a la construcción de un proyecto educativo", *Tecnología Educacional*, Associação Brasileira de Tecnologia Educacional (ABT), año XXI, nº 107, julio-agosto de 1992.

———: "Las configuraciones didácticas en la enseñanza universitaria: las narrativas metaanalíticas", *Revista del Instituto de Investigaciones en Ciencias de la Educación*, Facultad de Filosofía y Letras, nº 3, Buenos Aires, 1993.

———: "La evaluación de programas y proyectos: un viejo tema en un debate nuevo", en Puiggrós, Adriana y Krotsch, P. (comps.), *Universidad y evaluación. Estado del debate*, Buenos Aires, Rei Argentina, Aique, 1994.

———: *Las configuraciones didácticas. Una nueva agenda para la enseñanza superior*, Buenos Aires, Paidós, 1997.

Martí, Eduardo: "Metacognición: entre la fascinación y el desencanto", *Infancia y Aprendizaje*, nº 72, Madrid, 1995.

Mc. Cormick, Robert y James, Mary: *Evaluación del curriculum en los centros escolares*, Madrid, Morata, 1996.

Mortimore, Peter: "Autoevaluación escolar", en Galton, M. y Moon, B., *Cambiar la escuela, cambiar el curriculum*, Madrid, Martínez Roca, 1986.

Nisbet, John y Shucksmith, Janet: *Estrategias de aprendizaje*, Madrid, Santillana-Aula XXI, 1987.

Nunziati, Georgette: "Pour construire un dispositif d'évaluation formatrice", en *Cahiers Pedagogiques*, nº 280, 1990.

Ortega, Félix: "Unos profesionales en busca de su identidad", *Educación y Sociedad*, nº 11, Madrid, 1990.

Palou de Maté, María del Carmen: *La evaluación: propuestas para su análisis*, Cuadernos Universitarios, Universidad Nacional del Comahue, Facultad de Ciencias de la Educación, Cipolletti, Río Negro, 1992.

Popkewitz, Thomas: *Paradigma e ideología en investigación educativa*, Madrid, Mondadori, 1986.

Rosales, Carlos: *Evaluar es reflexionar sobre la enseñanza*, Madrid, Narcea, 1990.

Sánchez, Miguel Emilio: *Los textos expositivos. Estrategias para mejorar su comprensión*, Madrid, Santillana, Aula XXI, 1993.

Santos, Miguel Ángel: *Hacer visible lo cotidiano*, Madrid, AKAL, 1993.

Schön, Donald: *La formación de profesionales reflexivos*, Madrid, Paidós, 1992.

Selmes, Ian: *La mejora de las habilidades para el estudio*, Madrid, Paidós, 1988.

Stufflebeam, Daniel y Schinkfield, Antony: *Evaluación sistemática*, Barcelona, Paidós, 1987.

Tishman, Shari; Perkins, David y Jay, Eileen: *Un aula para pensar, aprender y enseñar en una cultura de pensamiento*, Buenos Aires, Aique, 1997.

Zemelman, Hugo: "Conocimiento y sujetos sociales", *Contribuciones al estudio del presente*, México, El colegio de México, 1987.

———: *Los horizontes de la razón. Dialéctica y apropiación del presente*, vol. I, Barcelona, Anthropos, 1992.

———: *Los horizontes de la razón. Historia y necesidad de utopía*, vol. II, Barcelona, Anthropos, 1992.

SISTEMAS DE CALIFICACIÓN Y REGÍMENES DE PROMOCIÓN*

Alicia R. W. de Camilloni

La calificación y la promoción de los estudiantes

Los sistemas de calificación empleados en la evaluación de los aprendizajes y los regímenes de promoción de los alumnos constituyen un componente esencial de los sistemas de evaluación que se emplean en la educación formal. Aunque ambos son producto de un conjunto de decisiones prácticas adoptadas en virtud de la adhesión a una u otra entre las diversas corrientes teóricas de la didáctica, siempre ejercen, de hecho, un fuerte efecto normativo sobre las acciones de enseñanza y de aprendizaje ya que, más que aparecer como una consecuencia de los fundamentos teóricos que se suscriben en la materia, tienen un efecto tan poderoso que suelen operar de manera muy efectiva, en reversa, como determinantes de todas las prácticas pedagógicas. De este modo, si los sistemas de calificación y promoción permanecen estáticos en sus modalidades fundamentales y no acompañan las transformaciones de las prácticas de la enseñanza, sólo sirven de obstáculo a esas transformaciones, obligando a los docentes a una permanente adaptación a los moldes pedagógicos que los procesos de renovación o de reforma pretenden reemplazar. Si, por el contrario, se resuelve producir modificaciones en los sistemas de calificación y promoción, éstos

* Este artículo ha sido posible gracias al subsidio otorgado a la investigación dirigida por la autora por la Secretaría de Ciencia y Técnica de la UBA (FI215).

tienden a desarrollar una indudable influencia sobre las estrategias de enseñanza de los docentes y, particularmente, sobre las estrategias de aprendizaje de los alumnos, con un efecto denominado de "validez retroactiva". Sobre estos procesos mucho se ha estudiado en el campo del análisis del denominado "currículo oculto". Aunque esta influencia puede ser abusiva, porque es necesario trabajar de manera integral sobre la enseñanza y la evaluación ya que ésta es subsidiaria de la primera, debemos reconocer que el papel de los sistemas de calificación y de los regímenes de promoción es central en la educación escolar, más allá de la pretendida desvalorización que de ellos se suele hacer en algunas teorías de la enseñanza, en tanto se les atribuye el constituirse en una manifestación no siempre adecuada o necesaria del ejercicio del control o de la autoridad interna o externa a la institución escolar sobre los aprendizajes de los alumnos y, en consecuencia, en una imposición que recorta la libertad personal de estos últimos.

La importancia que tienen, pues, los sistemas de calificación y los regímenes de promoción, sea evaluándolos en un sentido pretendidamente positivo o sea éste negativo, obliga a efectuar un estudio muy cuidadoso de las características que ofrecen las diversas modalidades que presenta cada uno de los modelos, de manera de seleccionar aquellos que resulten más apropiados para los propósitos de la programación didáctica, de la planificación de la institución y del sistema escolar.

Sistemas de calificación

Es preciso considerar, en primer lugar, dos cuestiones básicas. El primer aspecto relativo a la constitución de un sistema de calificación es el relacionado con la construcción y la fijación de una escala de calificación, y el segundo aspecto está referido a la decisión fundamental relativa a la adopción de sólo una o más de una de estas escalas en el sistema de calificación que se diseña. Veamos la primera cuestión.

En el proceso de construcción de una escala es necesario tener en cuenta las características que las escalas de calificación adquieren en términos de las condiciones y posibilidades que brinda cada una de las grandes modalidades de las escalas de medición en general.

Escalas de medición

Las escalas de medición pueden ser clasificadas en cuatro grupos a partir de las características que presentan respecto de la relación de ordenamiento progresivo u horizontal de las categorías que las constituyen, de la distancia igual o desigual que existe entre los diferentes grados de la escala y de la existencia o no de un punto de partida (cero absoluto) que permita operar de manera particular con tramos parciales de la escala.

De este modo, las escalas de medición se agrupan en los siguientes tipos:

a) *Escalas nominales*. En ellas las categorías que componen la escala no se encuentran en un orden de seriación que implique que un grado es mayor o mejor que el o los otros grados. No hay grados superiores e inferiores. Todos se encuentran en el mismo nivel. Se emplean para medir la frecuencia con la que los ejemplares evaluados se incluyen en cada una de las clases que la escala presenta. La construcción de una escala nominal requiere que se cumplan las siguientes condiciones: la escala debe incluir todas las categorías necesarias para diferenciar los ejemplares que se evalúan, dependiendo su carácter y número de la necesidad de obtener la información que se requiere con un fin determinado, aquel que dio lugar a la construcción de la escala. Para ello, las categorías se determinan en función del número indispensable de clases que permitan dar cuenta del universo que se estudia. El número debe ser el mínimo posible, sin que se pierda información útil. La escala, además, debe ser completa, es decir que debe hacer posible la inclusión en alguna categoría de la escala de todos los ejemplares analizados. Por otra parte, todas las categorías deben ser excluyentes entre sí, lo que implica que ningún ejemplar podría ser incluido simultáneamente en más de una categoría. La definición de las categorías, para ello, debe ser tan clara que no genere ninguna interpretación ambigua que pueda prestarse a error de clasificación. Una escala nominal se construye *ad hoc*, en los términos necesarios para el estudio que se desea realizar. Si se trata, por ejemplo, de clasificar a las personas según su nacionalidad, el número de categorías para incluir en la escala dependerá de la utilización que se quiera dar a las informa-

ciones que se recojan. En un caso puede ser suficiente contar con dos grados (Argentino; Extranjero) o con tres (Argentino nativo; Argentino naturalizado; Extranjero) o con cuatro (Argentino nativo; Argentino naturalizado; Natural de país limítrofe; Extranjero natural de país no limítrofe), etcétera.

b) *Escalas ordinales*. Son aquellas que presentan una serie de categorías ordenadas de acuerdo con un criterio definido por un sentido u orientación determinada que no permite introducir alteraciones en la secuencia. Los grados de la escala se ordenan, por ejemplo, de mayor a menor o de mejor a peor. Deben cumplir con los mismos requisitos que las escalas nominales, a los que se suma, como se ha dicho, la definición del criterio de ordenamiento de la serie. Se caracterizan, además, porque las distancias entre los pares de categorías o grados consecutivos no deben ser necesariamente iguales. Así, el intervalo entre "Muy bueno" y "Bueno" no tiene que ser igual al intervalo entre "Bueno" y "Regular". Las escalas ordinales son específicas y se usan para medir aspectos determinados. Las escalas ordinales no son y, en consecuencia, no deberían ser utilizadas como escalas universales. Sin embargo con cierta frecuencia adquieren una alto grado de formalismo y se usan indistintamente para evaluar cuestiones muy heterogéneas, perdiendo de este modo la riqueza posible de su valor descriptivo que deriva ya sea de su construcción *ad hoc*, ya de la especificidad de la definición de cada uno de los grados de la escala.

c) *Escalas de intervalos*. Son aquellas en las que se debe responder a los mismos requisitos que hemos enumerado para las escalas ordinales, con la particularidad que se agrega de que los intervalos entre pares de grados consecutivos deben ser iguales. Este último rasgo permite operar con las mediciones en las que se han utilizado las unidades medidas por estas escalas efectuando operaciones aritméticas que no pueden realizarse con las escalas ordinales. Entre ellas, por ejemplo, es posible obtener la media aritmética (promedio) de las mediciones, lo cual no tendría sentido con unidades de medición con escalas ordinales. Las escalas numéricas, siempre que los grados estén definidos sin ambigüedad de tal modo que aseguren la

igualdad de los intervalos, son un ejemplo de estas escalas intervalares.

d) *Escalas de razones o proporciones*. Son aquellas en las que no sólo se cumplen las condiciones establecidas para la construcción de una escala de intervalos sino que se agrega a éstas la de poseer un cero absoluto. Esta característica permite efectuar operaciones en las que se suman mediciones obtenidas independientemente o se comparan mediciones en las que, por ejemplo, el grado "2" de la escala se puede considerar equivalente a la mitad del grado "4". Si no existiera un cero absoluto, esta interpretación de la relación entre los grados 2 y 4 no tendría sentido.

Tipos de escalas de calificación

Para la construcción o selección de una escala y para las decisiones relativas a la adopción de una escala de calificación es indispensable tener en cuenta las características que hemos mencionado respecto de las escalas de medición en general, ya que ellas son la matriz básica a partir de la cual se pueden conceptualizar las escalas de calificación de los aprendizajes escolares.

La norma general que se ha de aplicar en la construcción de una escala de calificación es que el número de grados de la escala debe ser elegido de acuerdo con la posibilidad real que tiene el evaluador de operar con una técnica apropiada y disponible para él, discriminando los diferentes grados de calidad de aquello (rasgo, conducta, realización) que ha de ser evaluado con dicha escala.

Las escalas de calificación pueden presentarse en diversas formas. Las clasificaremos en escalas numéricas, escalas conceptuales o descriptivas y escalas gráficas.

La escala numérica consiste en cierto número de categorías establecidas en términos descriptivos a las que se atribuye de antemano valores numéricos convencionales.

El número de grados comprendidos por la escala puede variar. Con fines escolares se han utilizado escalas de muchos grados (100; 75; 60; 50; 25 o 20 grados) y escalas de pocos grados (11; 10; 6; 5; 4; 3 o 2 grados). En la escuela secundaria y en la educación superior argentina se ha utilizado tradicionalmente una escala de 11 grados

(0 a 10). De ella encontramos dos versiones. En una, la calificación mínima de aprobación es 4 puntos (con 4 grados o niveles de calidad diferentes para la desaprobación y 7 niveles de calidad para la aprobación). En la otra, la calificación mínima de aprobación es 6 o 7 puntos, con un número superior de grados para la desaprobación (de 0 a 5 o 6, respectivamente) que para la aprobación (6 a 10 o 7 a 10).

Se observa actualmente, sin embargo, una tendencia en la teoría de la evaluación a apoyar la propuesta de reducción del número de grados de las escalas de calificación a una cantidad que no supere los cinco grados. Este principio de reducción es fruto de las críticas provenientes de las corrientes anticuantitativistas en educación, las que no pueden aceptar aquello que, con toda evidencia, consideran un exceso en la diferenciación cuantitativa aplicada al análisis de las realizaciones de los alumnos. En la práctica se comprueba que, a pesar de la apariencia en contrario, las escalas integradas por un gran número de grados no permiten una discriminación rigurosa y clara de cada una de las categorías de calidad correspondientes. Por esta razón, lejos de perfeccionar la exactitud de la medición, contribuyen a acrecentar la imprecisión de las calificaciones asignadas por los docentes. También desde un punto de vista estrictamente práctico, trabajar con una escala de pocos grados facilita la tarea del docente ya que disminuye la posibilidad de error o injusticia y, por ende, la ansiedad en la elección de la calificación. Por otro lado, también ayuda a lograr consensos en los criterios de calificación empleados por diferentes docentes. De este modo, parece conveniente construir escalas con un número de categorías pequeño con el fin de mejorar la definición de cada grado y con el objetivo de contribuir al perfeccionamiento de la confiabilidad de las calificaciones.

Pero una escala de dos grados ("1" y "2" o "suficiente" e "insuficiente", por ejemplo) puede resultar poco descriptiva de los progresos de los estudiantes, porque cada grado debe cubrir un campo muy amplio y variado de rasgos. Si bien se aumenta la confiabilidad en las escalas de pocos grados, Robert E. Thorndike y Elizabeth P. Hagen (1991) señalan:

> Desde luego, esta ventaja se logra a expensas de la mayor parte de la información que el sistema de calificación puede suministrar acerca del individuo. En un sistema de aprobado-reprobado, solo sabemos que el

estudiante ha sido juzgado como aprobado en un curso, excepto por la pequeña fracción que se juzgó como fracasada. [...] A medida que se incrementa el número de discriminaciones más allá del sistema de dos o tres categorías, las distinciones empiezan a ser importantes para todos los estudiantes (pág. 577).

Así pues, si la capacidad descriptiva de la escala es pobre se perderán muchas de las funciones de la evaluación. Aumentando en pocos grados más la escala, con uno o dos puntos como mínimo por el lado de la no aprobación, y con dos o tres puntos por el costado de la aprobación, la escala permite una mejor definición de las categorías empleadas por el evaluador. Según los casos, sin embargo, puede ser útil discriminar con mayor fineza el nivel de logro en las realizaciones actuales de los alumnos. Las escalas de cuatro o cinco grados suelen servir de base para la elección de escalas. En ellas se emplean uno o dos grados para diferenciar la calidad en la desaprobación y dos, tres o cuatro para distinguir niveles de calidad en la aprobación. Por ejemplo:

1) 1 (desaprobación); 2, 3 y 4 (aprobación).

2) 1 y 2 (desaprobación); 3, 4 y 5 (aprobación).

Se puede afirmar que más allá del número que el análisis de los casos particulares demuestre ser necesario y suficiente, al aumentar el número de grados de la escala es más lo que se pierde en confiabilidad que lo que se gana en información.

Las escalas numéricas, según la calidad de la definición de cada uno de los grados que las componen, aplicando la categorización que antes se presentó con respecto a los tipos de escalas de medición, pueden ser ordinales, intervalares o de razones.

Las escalas conceptuales o descriptivas tienen algunos aspectos semejantes a las escalas numéricas pero presentan otros rasgos que los diferencian claramente de ellas. Ambos tipos presentan un escalonamiento de la calidad de lo que es evaluado desde un grado mínimo hasta un grado máximo. Se distinguen, empero, porque la escala numérica, al ser formalmente siempre la misma, sin importar el carácter sustantivo de lo que se evalúa, es mucho menos descrip-

tiva que la escala conceptual, se presta a distintas interpretaciones e induce a que rasgos heterogéneos sean tratados como equivalentes o equiparables. Una vez asignada una calificación, la frialdad de los números no permite discernir el tipo de rasgo, producto o comportamiento evaluado.

Las escalas conceptuales, por el contrario, se ajustan a las necesidades de cada instancia evaluativa. Varían su forma según las características de aquello que se desea evaluar.

Veamos algunos ejemplos:

Para apreciar la frecuencia con la que se produce un cierto fenómeno, la escala puede ser la siguiente:

- Siempre
- Generalmente
- A veces
- Nunca

Para apreciar la intensidad:

- Muy fuerte
- Fuerte
- Moderadamente fuerte
- Regular
- Débil
- Muy débil

Para apreciar la cantidad y la calidad de los conocimientos:

• Sobresaliente	• Excelente
• Distinguido	• Muy Bueno
• Bueno	• Bueno
• Aprobado	• Regular
• Insuficiente	• Insuficiente
• Reprobado	• Reprobado

Cada una de estas categorías constituye un grado de la escala y debe ser claramente definida. Con frecuencia se hace corresponder

a cada grado conceptual un símbolo numérico. Si esto ocurre, la escala conceptual se convierte en una escala numérica y presenta los mismos problemas de que adolece este último tipo de escalas. La delicadeza de las evaluaciones se pierde en favor de una formalidad presuntamente favorable a la exactitud, cómoda, por cierto, pero generadora del vaciamiento de la dimensión cualitativa de la evaluación.

En el mismo caso se encuentran las escalas alfabéticas como la que sigue:

A; B; C; D; E o F..., etcétera.

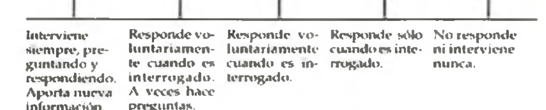
En ellas también suele establecerse una correspondencia con una escala numérica, pero esto no las convierte en escalas intervalares o de razones porque las escalas conceptuales son, por su misma naturaleza y básicamente, escalas ordinales.

La escala gráfica consiste en representar, sobre referencias lineales, categorías descriptivas determinadas.

Tiene la característica de presentar la intensidad de un rasgo como un continuo desde un grado mínimo hasta un grado máximo. La intensidad deseable del rasgo puede encontrarse señalada y definida en cualquiera de los puntos de la línea de evaluación, en ambos extremos y en un número determinado de puntos intermedios. Cada uno de los grados de la escala, cuyo número se recomienda que no supere los siete para no disminuir la confiabilidad de la evaluación, debe estar especificado en términos de lo que el docente está en condiciones de observar, conceptualizar e inferir realmente. En ninguno de los dos extremos se colocarán categorías que no puedan ser alcanzadas efectivamente o que se estima casi imposible que se presenten en los alumnos que integran la clase. Pero, en compensación, se habrá de tener en cuenta también que el apoyo gráfico lineal de la escala corresponde a una recta y no a un segmento. Será posible, por lo tanto, utilizar todos los puntos de la recta, aun los intermedios, entre los puntos definidos y los que se encuentran más allá de los últimos puntos definidos en los extremos.

Veamos algunos ejemplos de escalas gráficas.





Lecturas sobre la materia

Estudia a veces y superficialmente en el libro de texto.	Estudia siempre pero superficialmente en el libro de texto	Estudia siempre y cuidadosamente en el libro de texto.	Estudia en el libro de texto y consulta algunos materiales recomendados	Estudia en el libro de texto y consulta bibliografía recomendada y no recomendada por el docente.

Al igual que las escalas ordinales, las escalas gráficas se construyen *ad hoc*, no son escalas universales o formales sino escalas específicas.

Elección de una escala de calificación

Para determinar cuál ha de ser la escala que se ha de emplear en un programa de evaluación se deben tomar en consideración tres aspectos: el tipo de escala, el número de grados de la escala y la definición que se dará a cada uno de los grados.

En cuanto al número de grados, la escala debe estar compuesta por el número mínimo de categorías que resulte suficiente para que la evaluación no pierda sensibilidad ante las diferencias individuales. Un número razonable es, por lo general, de 4 o 5 grados, con uno

o dos grados para la desaprobación y tres para la aprobación. Si bien se han utilizado escalas de dos grados (aprobado-desaprobado), éstas parecen ser más adecuadas para los niveles más altos de la educación, por ejemplo, la enseñanza universitaria, pareciendo, sin embargo, insuficientes para diferenciar niveles de calidad en los trabajos realizados en los niveles primario o básico y secundario, en razón de que las calificaciones cumplen una doble función respecto del aprendizaje. En primer lugar, las escalas con diferencias entre diversos grados de calidad en la aprobación proveen información más precisa acerca de los resultados obtenidos y, en segundo lugar, influyen fuertemente, sin duda, sobre la motivación de los estudiantes frente a los requerimientos de estudio. El reconocimiento al esfuerzo realizado por el alumno aparece, en este caso, como un componente de la calificación y cuanto mayores son los niveles del sentimiento de éxito, mayor el nivel de aspiración del alumno. Ambos efectos, más y mejor información recibida por el alumno respecto de su aprendizaje y mayor motivación, actúan entre los determinantes más importantes del rendimiento.

Pero cuando las escalas tienen, en cambio, un número excesivo de grados, tanto por el lado de la desaprobación como por el de la aprobación, las escalas mismas pueden obligar a un exceso de cuantificación y convertirse en posible causa de error en la evaluación. El número de grados de la escala debe elegirse muy cuidadosamente. Requiere justificación explícita y la adopción de medidas complementarias que permitan un uso confiable de la escala elegida.

En cuanto a la definición de cada grado, debe ser clara y comprensible para todos los implicados en los procesos de evaluación y en el uso de la información que ella provee. Todos –docentes, alumnos, directivos, padres y público en general– deben poder interpretar las calificaciones de la misma manera. Cuando se emplea una escala numérica debe explicarse previamente qué significado se atribuye a cada uno de los símbolos numéricos que integran la escala. De otro modo, la evaluación estaría impedida de cumplir con sus funciones orientativas.

Se sugiere también la necesidad de completar la calificación numérica con descripciones que sirvan de guía a alumnos y docentes para el diagnóstico de logros y dificultades y para la orientación

del tratamiento apropiado para solucionar estas últimas. Una calificación numérica que no especificara en qué consisten los logros y los yerros no tendría utilidad para el alumno más que como pauta simbólica de su éxito o fracaso. No es de extrañar que con este tipo de calificaciones, la evaluación suela quedar al margen de los procesos de enseñanza y de aprendizaje y pierda, en consecuencia, sus funciones educativas. La calificación debe enriquecerse mediante la localización de los errores y el señalamiento aprobatorio de los resultados apropiados en el trabajo evaluado.

Otro difícil problema queda por resolver. Toda calificación lleva implícito un conflicto básico en las funciones de la evaluación. Hay dos maneras de calificar que responden a diferentes criterios, según se contrasten las realizaciones del estudiante con sus propias posibilidades o con un patrón ajeno a él, sea el rendimiento de sus compañeros (una norma) o una pauta establecida por el docente como objetivo a alcanzar (un estándar). Es evidente que cada una de estas formas dará lugar a calificaciones diversas para un mismo trabajo y un mismo alumno. También el éxito y el fracaso pueden tener dos tipos de interpretación. Tienen una realidad subjetiva cuando se refieren al acuerdo o desacuerdo existente entre el nivel de desempeño del alumno con su nivel de aspiración (lo que para un alumno es un éxito para otro puede ser un fracaso) y pueden tener una realidad objetiva cuando se compara ese nivel de desempeño con el patrón de rendimiento esperado y que ha sido adoptado a priori por el docente en su programación didáctica. Este conflicto de funciones se refleja, con frecuencia, en una pregunta común entre los docentes: ¿debemos calificar al alumno por su esfuerzo o debemos hacerlo en función de sus realizaciones efectivas? Ante esta pregunta se pueden ofrecer diversas respuestas porque, en efecto, el rendimiento del alumno puede ser apreciado, ya no de dos, sino de tres maneras diferentes, según se lo compare con los logros realmente alcanzados por otros alumnos, con una pauta fijada a priori por el docente o con el rendimiento que podría esperarse de él en condiciones conocidas.

Sin incursionar todavía en la problemática de los regímenes de promoción y restringiéndonos a la cuestión de las escalas de calificación, podemos establecer, siguiendo el primer criterio mencionado arriba, una relación entre el rendimiento de un alumno en una

asignatura determinada y el rendimiento medio de un grupo representativo de alumnos de edad similar en el mismo nivel, ciclo y año de enseñanza de la asignatura. Con este fin se han acuñado las nociones de edad educativa y de cociente educativo. De la misma forma en que se deriva el concepto de "edad mental" de un test de inteligencia, se deriva la edad educacional de un test de rendimiento. Utilizando un instrumento diseñado para evaluar conocimientos en un dominio determinado o para evaluar el desarrollo de una habilidad, se mide el rendimiento de una población de alumnos de características definidas. Cuando se trabaja con sujetos de distintas edades y niveles de aprendizaje en una disciplina, se obtiene, de este modo, una escala de normas de edad referida al instrumento utilizado. La edad educativa representa la edad en años y meses de las personas que obtuvieron un determinado puntaje en la administración del instrumento empleado, medido en términos de su relación con el puntaje promedio (en este caso, la mediana es la medida de tendencia central elegida) de la norma de edad de la población de referencia. La mediana es el puntaje por debajo del cual se encuentra la mitad de los puntajes, de todo ese grupo de edad. Con las medianas obtenidas se construye la escala de rendimiento por edad.

Tomemos, por ejemplo, la escala siguiente:

Puntaje	Edad	Año de escolaridad
20	14 años	8.8
21	14 años, 2 meses	8.9
22	14 años, 3 meses	8.9
23	14 años, 4 meses	9.0
24	14 años, 5 meses	9.1

Si un alumno obtuviera 22 puntos, su edad educativa sería de 14 años, 3 meses, la que correspondería a 8.9 años de escolaridad.

Si relacionamos la E.E. (edad educativa) con la edad cronológica (E.C.) del alumno obtendremos su cociente educativo (C.E.).

$$C.E. = \frac{E.E.}{E.C.} \ (.100)$$

Por ejemplo: $$C.E. = \frac{14 \text{ años, } 8 \text{ meses}}{13 \text{ años, } 8 \text{ meses}}$$

$$C.E. = \frac{176 \text{ meses}}{164 \text{ meses}} \ (.100)$$

$$C.E. = 107$$

El C.E. nos indicará cuánto tiempo ha necesitado el alumno para alcanzar un rendimiento determinado en comparación con el promedio (mediana) del tiempo requerido por los alumnos de su edad. Así, de la misma forma en que el C.I. se deriva de la noción de edad mental, el C.E. se deriva del concepto de E.E. En ambos casos se trata, y debe siempre puntualizarse, de un C.E. "medido", al igual que de una E.M. y un C.I. "medidos", sólo válidos con referencia a los instrumentos de medición utilizados.

Veamos algunos ejemplos. Un estudiante que tuviera 14 años de E.E. y una E.C. de 15 años, tendría un C.E. de 93.3. Otro estudiante de la misma clase que obtuviera el mismo puntaje en la evaluación pero que tuviera 13 años de edad cronológica, tendría un C.E. de 107. Se ha de tener en cuenta, empero, que para la determinación de la escala de normas de edad se toman como base las medianas de los puntajes respecto de cada grupo de edad y que, por lo tanto, en cada uno hay puntajes superiores y puntajes inferiores a las medianas. Al determinar los C.E. de los alumnos de una clase se ha de esperar, pues, que haya tantos estudiantes que estén por encima del valor 100 como por debajo de él. Hay que señalar, también, que el "C.E. medido" no se correlaciona directa y necesariamente con el "C.I. medido" pues son muchos los determinantes que actúan sobre el C.E. además del C.I.

Si tomamos, en cambio, el segundo criterio mencionado más arriba, que consiste en comparar al estudiante con una pauta establecida a priori por el docente, nos encontramos con la modalidad de la evaluación "referida a criterios". Aquí el verdadero problema a resolver es la determinación de los niveles de logro deseados por el docente, ya que ellos deben encontrarse entre los niveles posibles de alcanzar para los alumnos. El mecanismo que se utiliza para determinarlos es producto, por lo general y entre otros condicionantes, de la experiencia anterior del docente frente a sus alumnos, de la experiencia anterior del docente como alumno, de los niveles de aspiración del docente, de la cultura organizacional de la institución escolar en la que se desarrolla la enseñanza y, por último, de las tradiciones académicas de la disciplina que se enseña. Si las calificaciones del alumno son relativas a estos patrones, en ocasiones rígidos, y en otras oscilantes, la validez de dichas calificaciones puede ser cuestionable. Será siempre relativa a los criterios de base empleados y puede caer en un sensible irrealismo, tanto por lo bajo de la pauta establecida como por la imposibilidad de alcanzarla en razón de su extremada dificultad.

Si se sigue, por fin, el tercer criterio, consistente en comparar la realización del estudiante con lo que según sus condiciones podía esperarse de él, será necesario utilizar otra fórmula, que combine una evaluación de sus aptitudes con la evaluación de sus logros efectivos. Tradicionalmente para este fin se ha empleado el cociente de rendimiento (C.R.), obtenido dividiendo el C.E. por el C.I. Para un alumno que tiene, por ejemplo, un C.E. de 14 años de edad con una E.M. de 17 años, 6 meses, y con un rendimiento "normal" para su edad, el C.R. sería:

$$C.R. = \frac{C.E.}{C.I.} = \frac{\frac{E.E.}{E.C.}}{\frac{E.M.}{E.C.}} = \frac{\frac{14}{14}}{\frac{17{,}5}{14}} = \frac{100}{125} = 80$$

En este ejemplo, un alumno bien dotado no ha rendido todo lo que podía esperarse de él. La medida de esta diferencia nos la da el C.R. que es significativamente inferior a 100.

En otro ejemplo, un alumno de 15 años de E.C. y 13 años de E.M., retrasado en su E.E. (E.E.= 13 años) con respecto a la mayoría de los jóvenes de su misma E.C., tiene un C.R. de 100:

$$C.R. = \frac{C.E.}{C.I.} = \frac{\frac{E.E.}{E.C.}}{\frac{E.M.}{E.C.}} = \frac{\frac{13\ \text{años}}{15\ \text{años}}}{\frac{13\ \text{años}}{15\ \text{años}}} = \frac{86{,}6}{86{,}6} = 100$$

Este segundo alumno, cuyo cociente de rendimiento es 100, a pesar de sus dificultades, trabaja mejor que el primero en relación con lo que podía lograr cada uno de ellos. El primer alumno requiere mayor atención con el fin de ayudarle a alcanzar el pleno desenvolvimiento de sus posibilidades. Pero debemos tener siempre en cuenta, también, que se debe actuar sobre todas las potencialidades de los dos para no acentuar sus aparentes diferencias iniciales y lograr el máximo desarrollo, de hecho imprevisible, en ambos alumnos.

Los conceptos de C.E. y de C.R. se encuentran con mucha frecuencia en la literatura de habla inglesa sobre el tema y son, o eran, de aplicación muy frecuente en los países anglosajones. Su empleo ha sido escaso o nulo en América latina, en donde es también poco frecuente el empleo de pruebas estandarizadas de rendimiento. En los conceptos que hemos presentado pueden advertirse, por una parte, la influencia de las corrientes conductistas de la medición en la evaluación, que se manifiestan en el esfuerzo por generar fórmulas que permitan construir indicadores "exactos" y confiables para expresar los resultados de la medición. Pero, por otra parte, es interesante observar que, más allá del intento de cuantificar los rendimientos y las habilidades o aptitudes, estos conceptos permiten estudiar la relación existente entre aspectos que

en la evaluación de los aprendizajes con enfoques no conductistas también deben ser tomados en consideración. Son, pues, útiles, aun cuando no se pretenda utilizar las fórmulas para producir calificaciones numéricas y sólo se usen para contrastar criterios diferentes de evaluación. Su valor conceptual es significativo incluso en paradigmas no cuantitativistas de la evaluación en educación.

En estos términos, como hemos visto, la construcción de un sistema de calificación implica la elección entre alguno de los siguientes criterios: la calificación por las dotes que el alumno manifiesta poseer para la realización de tareas en dominios determinados, la calificación por sus logros efectivos en comparación con una pauta predefinida por el docente, o la calificación otorgada por comparación con los logros de los compañeros de su grupo de clase o de una población seleccionada. Cada uno marca una dirección distinta para la evaluación.

Al calificar al alumno con referencia a sus dotes y a su esfuerzo, se utiliza una noción correspondiente al campo del C.R. Un alumno voluntarioso, que logra superar sus propias dificultades, se vería mejor calificado que, por ejemplo, otro estudiante inteligente pero indolente. Las calificaciones de ambos estudiantes ya no serían semejantes en el caso de realizaciones iguales, sino que ocurriría –como con el C.R., si hubiera medio de calcularlo cuantitativamente– que el alumno esforzado obtendría mejor calificación.

Para calificar al alumno por su rendimiento efectivo en relación con una pauta predefinida, es menester expresar esta última en términos mensurables o apreciables para el docente. Se trata, entonces, de establecer previamente una descripción ajustada del rendimiento deseado. A este fin respondían los llamados "objetivos operacionales" que tenían como función principal facilitar una medición numérica ajustada. Esos objetivos pueden ser reemplazados, sin embargo, desde un marco teórico diferente, por la descripción cualitativa de los logros esperados. No se estaría lejos, en este caso, de la aplicación de la noción de C.E. porque, aun si no se desea trabajar con una cuantificación exacta, es obvio que el criterio que se aplica en una evaluación cualitativa puede coincidir con el de este concepto.

Cuando se adopta el criterio de calificar al alumno a partir de la comparación de sus aprendizajes con los logros efectivamente

demostrados por su grupo de clase u otro criterio tomado como referencia, los problemas que surgen son muy específicos y serán analizados en el examen que haremos posteriormente de los regímenes de promoción.

Aunque muchas veces ajenos al espíritu rigurosamente cuantitativista que sustenta el uso de calificaciones asignadas mediante sistemas asociados al uso de escalas de calificación numéricas de intervalos o razones, los docentes siempre comparan y eligen criterios. Es menester que se extremen los cuidados para que las calificaciones tengan origen y fundamentación clara, en particular, para los alumnos.

Considerando que, entre los criterios que hemos expuesto más arriba, el equivalente al C.E. y el semejante al C.R. evalúan en función de aspectos diferentes, ambos interesantes por su significatividad, se puede sostener que sería conveniente utilizarlos en conjunto, con el propósito de mejorar las apreciaciones parciales y de no perder información útil para el docente y el alumno. El sistema que se crea de este modo debiera conservar la identidad de cada uno y preservarlos de la confusión en la interpretación. Una sola calificación numérica, apoyada en uno solo de esos criterios o como fruto de una combinación de ambos, sería siempre insuficiente para el cumplimiento de las funciones de la evaluación. Se recomienda adoptar, entonces, un sistema de doble calificación. En él se complementan dos calificaciones: una, referida a los progresos realizados por el escolar en relación con una norma establecida previamente, y la otra, referida a sus progresos contrastados con sus aptitudes consideradas individualmente. Este sistema permite combinar las ventajas de ambos criterios y permite manejar simultáneamente los dos sentidos, objetivo y subjetivo. Para realizaciones semejantes, por ejemplo, las dobles calificaciones podrían resultar las siguientes:

Alumno A: 4 (muy bueno)
Alumna B: 4 (regular)
Alumno C: 4 (excelente)

Aquí se han empleado dos escalas: para el rendimiento efectivo en términos de una pauta predefinida: escala numérica de 1 a 5. Para

el rendimiento en términos de capacidad y esfuerzo, escala conceptual con grados con definición clara: malo, regular, bueno, muy bueno, excelente.

La primera calificación procura ser objetiva. La segunda, más subjetiva, se puede emplear en la construcción de una calificación nueva atendiendo al sentido de la progresión que se manifiesta en estas segundas calificaciones. Servirá en cada caso de advertencia o de estímulo y reconocimiento del docente ante la actuación del alumno, incidiendo sobre posteriores aprendizajes y, por tanto, como proveedora de información y estímulo de la motivación. Las segundas calificaciones pueden convertirse en calificación numérica. Pero ésta no resulta del cálculo del promedio sino de la tendencia que ha manifestado el alumno respecto de su esfuerzo en el trabajo de aprendizaje.

La necesidad de completar las evaluaciones numéricas con anotaciones referentes a los errores y sus posibles causas, a la adecuación de la velocidad impresa a la realización del trabajo, que puede ser excesivamente lenta o apresurada, y la referencia a la dirección e intensidad de los esfuerzos del alumno, nos muestra la insuficiencia de un sistema de calificaciones que pretenda manejarse sólo con una simple escala numérica.

La enseñanza opera como una fuerza en un sistema de fuerzas. Es por sí misma, ella también, un conjunto dinámico que actúa sobre y con sujetos dinámicos. Las magnitudes escalares, desprovistas de orientación y en las que sólo interesan los valores numéricos, no pueden describir el tipo de cambio producido por la educación. Por esto, pensamos, el aprendizaje escolar debiera ser conceptualizado a los fines de su medición como magnitud vectorial en el marco de la evaluación, representándose los progresos por segmentos o semirrectas con dirección, y no sólo por sus valores numéricos asignados en instancias puntuales aisladas entre sí.

Decisión sobre el número de escalas de calificación

Del análisis efectuado sobre los distintos tipos de escala de medición y de calificación surge con claridad que cada una de las modalidades califica objetos (caracteres y rasgos, procesos y productos) diversos y lo hace de distinta manera. De ahí que toda

calificación sea relativa no sólo al instrumento que se ha utilizado para medir, como lo afirmamos más arriba, sino también a la escala de calificación empleada.

En la mayor parte de los sistemas se ha intentado construir escalas de carácter universal, esto es, capaces de describir con precisión aceptable los diferentes niveles de calidad de objetos, procesos o habilidades y destrezas de todo tipo. Estas escalas tienen carácter formal y pueden ser ejemplificadas por las escalas numéricas que son de uso habitual en la educación escolar. Cuando las escalas alfabéticas o conceptuales están integradas por grados definidos de manera muy general, también se emplean al modo de escalas universales. Este tipo de escalas presentan la obvia ventaja de su practicidad. Todos los docentes pueden usar la misma escala y se hacen posibles, también, las comparaciones entre calificaciones correspondientes a distintas disciplinas y clases de trabajo. Se produce, de este modo, una homologación formal de la consistencia interna de los logros de un alumno y, también, de los de una cantidad ilimitada de alumnos. Pero se debe remarcar que la homologación es sólo formal y que, por esta causa, pierde la calificación buena parte de su valor descriptivo. Si se quieren registrar y conservar las características diferenciales específicas de las realizaciones o de los aprendizajes de los alumnos, tratándose de uno solo o de muchos, es imprescindible emplear escalas diferentes con menor pretensión de universalidad, ya que ésta se encuentra en correlación inversa con el poder descriptivo de las calificaciones. La decisión de emplear una misma escala, por ejemplo, para todas las disciplinas del currículo o de emplear escalas construidas *ad hoc* para cada una de ellas es, pues, una cuestión de la mayor importancia. De lo expresado hasta ahora al respecto, se infiere que parecería que el problema de las escalas de calificación en un programa escolar estará bien resuelto cuando esté conformado por el menor número de escalas diversas que resulten necesarias para proporcionar la información que sea útil para los alumnos, los docentes, las autoridades escolares y todos aquellos implicados en los procesos de enseñar y aprender en la escuela.

Regímenes de promoción

El establecimiento de un régimen de promoción es siempre una tarea delicada en la que los riesgos de adoptar decisiones equivocadas son grandes. El régimen elegido debiera hacer imposible que se impida la promoción de un alumno al curso siguiente si éste se encuentra en condiciones de proseguir sus estudios porque ha realizado ya los aprendizajes necesarios. E, igualmente, tampoco debiera permitir la promoción de un alumno que no posee los conocimientos básicos indispensables para realizar los aprendizajes correspondientes al curso superior o que no ha alcanzado las competencias que la acción escolar debe garantizar ante el propio individuo y la sociedad. La generosidad injustificada en materia de promociones es causa muy frecuente de posteriores fracasos escolares, así como el a veces injusto rechazo de promociones se produce, en ocasiones, en el sistema escolar y es causa, a su vez, de grave frustración personal, e incluso de abandonos y deserción escolar. Numerosos estudios comprueban la importancia que tienen las calificaciones obtenidas en la escuela sobre la autoestima de los alumnos. La influencia de las calificaciones es grande y tiene un doble efecto: por un lado y de manera general, sobre la construcción del "sí mismo" del estudiante, y por el otro, y más específicamente, sobre las aspiraciones que tiene frente al estudio. Si el primero es decisivo en el proceso de desarrollo de su personalidad, el segundo tiene un efecto primario en las motivaciones y los niveles de rendimiento deseados y efectivamente alcanzados por él.

La elección de un régimen de promoción presenta, entre sus principales dificultades, la de requerir que se asocien de manera previsible las calificaciones provenientes de diferentes instancias, lo cual implica encontrar una forma de conjunción de lo diverso. Los términos "agregación", "combinación" y "reconciliación" han sido adoptados para referirse a los procesos donde se reúnen y relacionan conjuntos de calificaciones con el propósito de producir una calificación única. Aunque afirma que no es posible producir definiciones perfectas de estos términos, Dylan Wiliam (1992) propone utilizar la palabra "combinación" para describir todos esos procesos, "agregación" cuando se reúnen evaluaciones

de diferentes "cosas" y "reconciliación" cuando las calificaciones se refieren a la misma "cosa" (pág. 19).

Para la fijación de un régimen de promoción son varios los criterios que se han de combinar o complementar, luego de analizar los rasgos fundamentales de las modalidades empleadas con mayor frecuencia en los sistemas escolares.

Analizaremos, en primer lugar, seis de ellos, que se destacan por su uso más frecuente en los sistemas escolares: el sistema de promoción por calificación promedio, el sistema de "logros mínimos exigidos", el sistema de promoción por examen final, el así llamado "sistema de promoción sin examen", los sistemas de puntajes derivados y la evaluación de portafolio.

Régimen de promoción por calificación promedio

Emplea una de las medidas estadísticas de tendencia central como base para la determinación de la promoción. La medida más utilizada con este fin es la media aritmética. Ésta se obtiene sumando primero todas las calificaciones obtenidas por el alumno y dividiendo después la suma por el número de calificaciones sumadas:

$$\text{Media aritmética} = \frac{\text{Suma de x}}{\text{n}}$$

en donde *x* son las calificaciones numéricas obtenidas por el alumno y *n* el número de calificaciones independientes sumadas.

Las calificaciones de bajo puntaje son equilibradas por calificaciones de puntaje más alto. En el régimen de promoción se determina el promedio mínimo que debe alcanzar el alumno para obtener esa promoción.

Este régimen de promoción tiene algunas características que deben tenerse en cuenta cuando se elige un sistema de promoción o cuando se aplica como resultado de la selección efectuada por la institución o el sistema escolar.

Para que las calificaciones promedio sean confiables debe cumplirse la condición de que la escala empleada sea una escala de intervalos o una escala de razones. Si los intervalos entre pares

de grados consecutivos no son iguales no tiene sentido la pretensión de realizar operaciones aritméticas que pretenden dar cuenta exacta de resultados de aprendizaje traducidos en números que expresan tanto logros como diferencias entre alumnos con una precisión no justificada. Si en una escala, por ejemplo, la distancia entre dos grados consecutivos (punto 5 y punto 6 de la escala) no es igual a la distancia que existe entre otros dos (punto 6 y punto 7), el promedio de las calificaciones 5 y 7 no puede ser 6 puntos ya que la distancia entre 5 y 6 no puede ser compensada por la distancia 6 a 7 puntos. Ahora bien, en la práctica, en las escalas con calificación promedio mínima exigida para la aprobación es común que la distancia entre esa calificación y la inmediata inferior sea mayor que la que existe entre el promedio mínimo exigido y la calificación inmediata superior a él. El docente suele exigir mayor rendimiento relativo para pasar de la calificación inmediata inferior al promedio mínimo que se exige para promover al alumno, por ejemplo, el que el mismo docente requiere para asignar la calificación inmediata superior a éste. Por lo tanto, la escala utilizada en la práctica, a pesar de su apariencia, no es una escala de intervalos o de razones, porque se constituye en ella una zona límite, de extensión variable y definición ambigua, en torno a la calificación promedio y en la que, con frecuencia, puede no haber diferencia efectiva entre los alumnos promovidos y los no promovidos.

Compárense las calificaciones siguientes que corresponden a diferentes alumnos en un régimen que exige una calificación promedio mínima de 6 puntos para la aprobación:

A: 4; 6; 6; 7 .... calificación promedio: 5,75
B: 7; 6; 7; 4 .... calificación promedio: 6
C: 5; 7; 6; 6 .... calificación promedio: 6
D: 7; 6; 5; 7 .... calificación promedio: 6,25

Las calificaciones promedio, en este caso, muestran cómo se construye esa zona límite. Para salvar el concepto de zona, en reemplazo del punto límite para la promoción (que haría que el alumno A no fuera promovido y sí los tres restantes), es común que se aproximen la calificaciones que se encuentran entre n,50 y n,99 al entero más cercano. Pero este esfuerzo de corrección de la escala no

es suficiente para evitar y justificar la inexactitud real de la calificación. También habría que demostrar que existe real diversidad entre el n,49 y el n,50. Por otra parte el sistema tampoco permite que se manifiesten con claridad en los promedios las diferencias de rendimiento entre los cuatro alumnos en las diferentes instancias de evaluación.

Como vemos, la calificación promedio no es suficientemente descriptiva del nivel de aprendizaje de los alumnos. Esto es producto de la conjunción de diversos factores. Entre ellos podemos mencionar: 1) la variación de las distancias entre grados consecutivos en las escalas de calificación que se utilizan habitualmente en la escuela, y 2) la combinación de este sistema de promoción con escalas que poseen un número relativamente grande de grados. Si se emplea, por ejemplo, una escala de 11 grados (0 a 10) con el sistema de promoción por promedio calculando hasta los centésimos, la escala resultante tendrá 1001 grados. Si la escala de base tuviera 5 grados, la escala para los promedios tendría 401 grados.

Como ya vimos, esta excesiva discriminación cuantitativa entre valores asignados a conjuntos de aprendizajes no es válidamente descriptiva de éstos ni de las diferencias existentes entre los aprendizajes de distintos alumnos. Sería difícil demostrar que la diferencia entre 574 y 575 puntos tiene alguna correspondencia significativa cuando se interpretan logros y fracasos en esos aprendizajes. Del mismo modo no puede considerarse relevante una diferencia de 25 puntos en una escala de 1001 grados.

El sistema de promoción por promedio presenta otras serias desventajas:

1. Un alumno que ha hecho un trabajo excelente y uno muy malo obtendrá la misma calificación promedio que otro que ha hecho dos trabajos pobres. Las calificaciones de uno y otro, por ser idénticas, aunque se refieren a dos situaciones sustancialmente distintas, no son útiles porque no pueden servir de base ni para una tarea de diagnóstico ni de pronóstico.

2. Por este sistema de promoción adquieren el mismo valor, respecto de la calificación promedio, trabajos de naturaleza heterogénea: orales, escritos, prácticos, teóricos, breves, extensos, menos importantes, muy importantes, etcétera.

3. Marca un tope de rendimiento para los alumnos. La calificación promedio mínima exigida para la promoción se convierte en el objetivo perseguido y en el símbolo del nivel de aspiración de una gran cantidad de alumnos, incluso de muchos estudiantes que podrían alcanzar mejores logros.

4. Crea en el docente una tendencia a cuantificar todos los comportamientos de los alumnos en clase, estereotipando las categorías de observación.

5. Se presta a las especulaciones de los alumnos con respecto a la calidad de los trabajos que presentan, funcionalizándolos en relación con su necesidad de lograr sólo el puntaje necesario para obtener el promedio requerido para la promoción.

6. El régimen de promoción por promedio facilita que la ignorancia en un tema o el mal desempeño en una tarea sea compensado por el conocimiento de otro tema o la realización de otra tarea diferente.

Algunas de estas características no deseadas son propias de la naturaleza misma del sistema, otras, en cambio, nacen de su mala utilización cuando no se adoptan los recaudos debidos. En los dos casos afectan seriamente el aprendizaje de los alumnos. No se puede negar, sin embargo, que su amplia aplicación se debe a su practicidad, porque facilita el manejo de gran cantidad de calificaciones y de decisiones de promoción o no promoción de los alumnos.

Se puede intentar reemplazar la promoción por promedio (media aritmética) por algún otro sistema que emplee también una medida de tendencia central diferente, la mediana o el modo.

La mediana presenta la desventaja de no tomar en cuenta los puntajes extremos. Si un alumno tuviera las siguientes calificaciones: 5; 6; 7; 8; 9, la mediana sería 7 (puntaje debajo del cual cae el 50 % de las calificaciones). Pero seguiría siendo 7 si el 5 fuera reemplazado por un 1 y el 9 por un 10. No es, por lo tanto, más representativa del nivel de aprendizaje del alumno que la media aritmética. En cuanto al modo, que es el puntaje que se presenta con mayor frecuencia en una serie de calificaciones de un alumno, sólo sería realmente representativo si el estudiante tuviera un número grande de calificaciones, lo cual no es muy habitual. Es corriente, por otra parte, que si la escala tiene un número de grados superior a cuatro

o cinco, del procesamiento de los puntajes de un alumno se deriven dos o más modos. Podemos afirmar, sin embargo, que si se considerara adecuado utilizar una medida de tendencia central, se podría interpretar que el modo es más descriptivo del nivel de aprendizaje típico alcanzado por el alumno que la media aritmética o la mediana.

El problema parece ser que es necesario encontrar un sistema que permita tomar en cuenta todas las calificaciones del alumno, incluyendo los que se encuentran alejados del punto medio de la distribución, ya que ellos reflejan, también con claridad, los déficit y las posibilidades del estudiante.

Un régimen que permite paliar algunas de las deficiencias del promedio (media aritmética) y que conserva la practicidad de éste, permitiendo, a su vez, tomar en cuenta todos los puntajes obtenidos, es la media ponderada. Para calcularla se establece previamente una categorización de las producciones del alumno, según su carácter, su extensión y su importancia, y luego, una ponderación numérica para cada tipo de trabajo o instancia de evaluación. Esta categorización y su ponderación deben ser comunicadas a los alumnos. A la hora de calcular el promedio (media aritmética), cada calificación se multiplica por la ponderación correspondiente. Se suman todos estos resultados y a continuación se divide la suma total por la suma de las ponderaciones empleadas. Si, por ejemplo, los trabajos han sido clasificados por su importancia en tres tipos: los muy importantes, los medianamente importantes y los importantes (ya que suponemos que si no fueran de algún grado efectivo de importancia no serían tomados en cuenta) y se hubieran asignado respectivamente las siguientes ponderaciones: 3; 2 y 1, entonces, el cálculo de la media aritmética ponderada se efectuaría como sigue:

Calificaciones obtenidas por un alumno X

Categoría "muy importantes":	9 (3); 8 (3); 9 (3)
Categoría "medianamente importantes":	5 (2); 8 (2)
Categoría "importantes":	4 (1); 6 (1)

La media ponderada se calcula del siguiente modo:

$$\frac{9\,(3) + 8\,(3) + 9\,(3) + 5\,(2) + 8\,(2) + 4\,(1) + 6\,(1)}{3 + 3 + 3 + 2 + 2 + 1 + 1}$$

$$= \frac{27 + 24 + 27 + 10 + 16 + 4 + 6}{15} =$$

$$= \frac{114}{15} =$$

$$= 7{,}60 \text{ puntos}$$

Si se calculara una media aritmética no ponderada, la calificación promedio sería de 7 puntos. Como se ve en este caso, el mayor peso de las instancias "muy importantes" mejora la calificación final.

Se trata en los casos que hemos presentado hasta aquí, de calificaciones puntuales, obtenidas en instancias separadas una de la otra, pero se puede adoptar otro punto de vista respecto de la evaluación final de un curso. De Ketele (1984) atribuye a la evaluación, entre otras, la función de certificar que la persona evaluada posee los conocimientos suficientes para pasar al próximo año, al curso siguiente o a la vida profesional, según proceda. Pero, en esos casos, señala la importancia de que se tomen en cuenta dos aspectos: 1) "este tipo de evaluación se refiere a los objetivos terminales, es decir, a los macroobjetivos que integran un número relativamente importante de objetivos intermedios o microobjetivos, así pues, la evaluación terminal no se refiere a estos últimos. Nos atreveríamos a decir –afirma De Ketele– que en cierta manera pueden olvidarse como tales, es su integración lo que nos interesa"; y 2) "la certificación es una decisión dicotómica: el alumno tiene o no la competencia mínima; interesa, pues, definir bien los criterios" (pág. 30). Así, de modo consistente, en el modelo de la Pedagogía de la Integración que propone De Ketele (1993), la evaluación certificativa final toma como objeto para evaluar el objetivo terminal de integración, en tanto los objetivos intermedios son evaluados por evaluaciones formativas o diagnósticas (pág. 70). En atención a estas observacio-

nes, por ende, conviene emplear otro régimen de promoción, diferente de la calificación promedio. Pero, conviene señalar, que si éste último fuera el régimen establecido en la institución o en el sistema escolar, una adaptación que lo mejoraría consistiría en la introducción del concepto de una calificación promedio ponderada (media ponderada) en la que la ponderación se estableciera diferenciando el peso de las calificaciones que corresponden a los que De Ketele denomina microobjetivos, con un coeficiente de valor 1, por ejemplo, y macroobjetivos con un coeficiente de valor 2 o 3.

Régimen de promoción por logros mínimos exigidos

Este régimen se establece sobre la base de la exigencia de un cierto rendimiento mínimo preestablecido en todas las competencias determinadas como significativas para la promoción. A diferencia del régimen de promoción por promedios, la decisión de promoción no resulta del cálculo de una sumatoria ni de una síntesis que globaliza rendimientos heterogéneos sino del análisis del cumplimiento de cada tarea considerada por sí, en tanto indicadora del logro de la o las competencias correspondientes, las cuales no pueden ser sustituidas por otras competencias distintas. El propósito es garantizar que el estudiante haya alcanzado todos los saberes que se consideran necesarios. Los niveles mínimos de calidad de la tarea para su aprobación son fijados previamente, por lo general, por el docente o por la escuela. Este régimen tiene algunas ventajas sobre el de promoción por promedios ya que asegura un mejor trabajo del alumno a lo largo de todo el curso y hace inútiles las especulaciones con la calificación que son típicas de ese régimen. Pero presenta, también, algunos riesgos que es imprescindible tener en cuenta para evitar su efecto perjudicial. Entre ellos, tiende, al igual que la promoción por promedio, a fijar un tope muy bajo de rendimiento, al nivel del que se ha fijado como rendimiento mínimo aceptable. Si en una escala 0-10, el mínimo de aprobación se estableciera en 5 puntos, éste se convertiría en el objetivo de aprendizaje de muchos alumnos. Los grados superiores de la escala podrían ser percibidos como superávit no necesarios por algunos estudiantes.

Existe siempre, por otra parte, el riesgo de que un alumno que hubiera fracasado en una de las tareas iniciales del curso, en el

entendimiento de que la promoción para él sería imposible o, al menos, muy difícil de alcanzar, perdiera interés y abandonara ese curso sin hacer el esfuerzo necesario para remontar los resultados negativos del comienzo.

Estos problemas, en un régimen que tiene ventajas, pueden ser obviados hasta cierto punto por un buen manejo del docente. Se debería mostrar la importancia de los aprendizajes profundos frente a los superficiales que el régimen podría estar estimulando y ofrecer oportunidades reales de recuperación en relación con tareas en las que el alumno hubiera mostrado tener dificultades, eliminando el carácter definitivo de las desaprobaciones iniciales.

La determinación de los estándares mínimos de aprendizaje para la aprobación es una labor de gran importancia y complejidad. En primer lugar, los estándares deben estar definidos con la mayor claridad posible. Además, no deben ser demasiado bajos porque anularían la eficacia del sistema. Pero tampoco excesivamente altos porque se convertirían en exigencias imposibles de alcanzar. En ambos casos, de excesiva facilidad o excesiva dificultad, los niveles de aspiración de los alumnos sufrirían un descenso, generando un nivel correlativo menor de calidad del aprendizaje. Por lo demás, y desde un punto de vista práctico, la implementación del régimen por logros mínimos no presenta mayores dificultades siempre que se compensen los peligros mencionados con estímulo adecuado para que cada alumno alcance su mejor rendimiento de acuerdo con sus posibilidades y se evite la fragmentación de los aprendizajes en objetivos parciales.

Régimen de promoción por examen final

En este régimen se requiere que, para obtener su promoción, el estudiante apruebe una instancia de evaluación escrita, oral o de actuación, o una combinación de dos o de todas ellas. El examen debe estar diseñado de manera tal que permita apreciar en síntesis el aprendizaje logrado por el alumno a lo largo de todo el curso.

No puede tratarse de una simple prueba de actuación o ejecución de una tarea, sino que debe tener un alto grado de validez de contenido y de validez predictiva. Debe representar, en consecuencia, una muestra significativa del universo de contenido de la

asignatura, a la vez que debe predecir eficazmente el desempeño posterior del alumno, sea dentro o fuera del ambiente escolar.

Puede utilizarse de diferentes maneras:

- Como criterio único para decidir la promoción de los alumnos, en cuyo caso el trabajo a lo largo del curso tiene como un objetivo principal preparar al estudiante para la aprobación de la instancia de evaluación final. Las calificaciones parciales, si existen, sirven de orientación al aprendizaje y a la enseñanza y, a veces, para seleccionar a los alumnos a los que se permite rendir el examen final con probabilidad de éxito.
- Combinado con otras formas de promoción, el examen final provee calificaciones que se computan en la calificación que decide la promoción. Un ejemplo es el cómputo del puntaje obtenido en este examen como parte de una calificación promedio, ponderada o no.
- Como supletorio de otras formas de promoción cuando el alumno no ha obtenido ésta por medio del sistema regular establecido en primer término.

El examen final o instancia evaluativa equivalente, presenta ventajas y también desventajas tanto cuando se emplea como régimen único como cuando sólo tiene una participación parcial en la decisión de promoción.

Algunas desventajas se originan en el uso del examen como técnica de evaluación y otras en la adopción del examen como criterio definitivo para la promoción. En lo que respecta a la validez predictiva de los exámenes, ésta requiere que permita un pronóstico de las competencias de los alumnos en un período extenso y en situaciones escolares y extraescolares. Las pautas utilizadas como criterio de evaluación deben permitir la elaboración de inferencias de alto grado de probabilidad acerca de su desempeño futuro. "Como regla –decía Kurt Lewin– cuanto más macroscópica sea la situación que tenga que ser descrita, más largo será el período que habrá que observar para determinar la dirección y la velocidad de la conducta en un tiempo dado."*

* Citado por Ernest R. Hilgard: *Teorías del aprendizaje*, México, FCE, 1961, pág. 161.

La permanencia y desarrollo de las competencias que deben ser pronosticadas en los exámenes escolares se extienden de modo muy prolongado en el tiempo futuro ya que la escuela tiene fines de largo alcance. El examen, por tanto, de una extensión temporal acotada, no puede ser utilizado como único criterio ya que describe sólo la actuación del alumno en una situación de breve duración. Las hipótesis acerca de las ejecuciones futuras del alumno están aquí en dependencia de su rendimiento en un período de duración muy limitada y están viciadas por la emergencia de factores circunstanciales. Por esta razón no debiera utilizarse de modo exclusivo para decidir acerca de cuestiones fundamentales relativas a la selección o promoción de los alumnos. Su validez y confiabilidad son mayores cuando se emplea complementariamente a otros instrumentos de evaluación o de manera supletoria, como antes se dijo.

El sistema de promoción que incluye una instancia de evaluación final tiene un valor, es menester señalarlo, que lo hace irreemplazable en un programa de evaluación, ya que permite evaluar de manera completa, no parcelada, el dominio alcanzado por el alumno sobre el universo de los contenidos de la asignatura, el área o un conjunto de asignaturas. En los sistemas de promoción por promedio y de logros mínimos exigidos se emplean evaluaciones parciales, por lo que en ellos queda siempre el riesgo de que falte la apreciación global del aprendizaje del universo de contenidos que se debe tomar como unidad. Y, como se ha señalado más arriba, el problema principal no es la cuestión de que, en el marco de la evaluación considerada aisladamente, se reduzca el grado de validez y confiabilidad de las calificaciones asignadas, sino la influencia que ello tiene sobre el aprendizaje que, consecuentemente, hacen y harán los alumnos. El efecto retroactivo de la modalidad de la evaluación elegida es negativo en este caso, porque induce a los alumnos a un estudio que puede ser muy serio y riguroso en un enfoque analítico pero pobre desde la perspectiva del establecimiento de relaciones entre las partes y del desarrollo de la capacidad de construir una visión integral de los contenidos estudiados.

Régimen de "promoción sin examen"

La modalidad de régimen de promoción que suele denominarse "promoción sin examen" se define por su diferenciación respecto del empleo de un examen final. No es un régimen del que se pueda ofrecer una definición unívoca. Los programas de evaluación en los que se expresa pueden asumir diversas formas, tratándose, por lo general, de la aplicación de un régimen de evaluaciones parciales cuyas calificaciones se procesan según el régimen de promoción por calificación promedio o de logros mínimos exigidos. En muchos casos se incluye una instancia de evaluación final integradora que ocupa el lugar de una evaluación parcial más, como las otras anteriores y, a la vez, de examen final por el papel que desempeña en el proceso de aprendizaje del alumno en la fase de preparación para enfrentar esta instancia que hace las veces de un examen final complementario.

Una modalidad peculiar que puede adoptar es la que corresponde a un programa de evaluación continua basado sobre el análisis, interpretación y apreciación de los registros sistemáticos de las observaciones efectuadas por el docente respecto de las conductas, los procesos y las producciones realizadas por los alumnos en situaciones no puntuales de evaluación. Estas últimas son situaciones de trabajo no diferenciadas de la tarea de aprendizaje. Orientación y guía del aprendizaje y evaluación del aprendizaje son tareas que el docente realiza de manera simultánea. La evaluación es continua y no se diferencian las instancias o momentos destinados a la evaluación de aquellos en los que el docente "enseña". Se trata de un trabajo de carácter clínico, muy individualizado, que se puede llevar a cabo cuando se trabaja con grupos pequeños y en el que la autoevaluación del alumno juega un papel de gran importancia, ya que no sólo es un instrumento útil de evaluación sino una meta de la enseñanza por su valor intrínseco en la formación del alumno.

Régimen de promoción por puntajes derivados

Los puntajes derivados son normas resultantes del procesamiento estadístico de los puntajes brutos obtenidos efectivamente por los alumnos del grupo que se toma como referencia. Los grupos son,

usualmente, el grupo de clase a la que pertenece el alumno, su cohorte, su grupo de edad o el grado o año de escolaridad que corresponde. El procesamiento de los puntajes obtenidos por los estudiantes cuando les son administradas las pruebas se realiza mediante la aplicación de reglas fijas. Las normas no representan niveles deseables de rendimiento definidos apriorísticamente, sino construcciones de los evaluadores que indican los niveles que se ha demostrado que los alumnos pueden alcanzar de manera efectiva. Se procura corregir, de este modo, el nivel de logro establecido en los objetivos de aprendizaje propuestos a priori por el docente, la institución o el sistema escolar para adecuarlos a las posibilidades reales de aprendizaje que tuvieron los alumnos y que pudieron demostrar en las pruebas de evaluación como resultado de la enseñanza que recibieron. De este modo, los puntajes derivados se diferencian de los que responden a la aplicación de la llamada "evaluación referida a criterios" que emplea, en cambio, escalas que tienen un valor absoluto y que corresponden a los niveles de logro determinados en objetivos previamente establecidos definidos en términos de aprendizajes deseables. Ambos enfoques, si bien diferentes uno del otro, responden por igual a la búsqueda de la construcción de procedimientos algorítmicos para la asignación de calificaciones a los alumnos, de acuerdo con los principios que han sido desarrollados por el programa de investigación científica de la psicología conductista y sus derivaciones didácticas. W. James Popham (1983) señala la importancia y la autoría de esta distinción:

> [...] el mérito de haber creado una estrategia de solución le corresponde a Robert Glaser, que en 1963 publicó un ensayo que ya se considera clásico en el cual introdujo las expresiones evaluación basada en pautas de normalidad y evaluación basada en criterios. Glaser empleó el concepto de evaluación basada en normas estadísticas para describir los tradicionales tests de rendimiento en los cuales se pone el acento en la determinación de la posición relativa de un sujeto dentro de un grupo. Por otra parte, aplicó el concepto de evaluación basada en criterios a los tests que identificaban el dominio (o ausencia de dominio) de "conductas específicas" por parte del sujeto examinado (pág. 31).

Igualmente Norman E. Gronlund (1973) define a estas pruebas indicando que el criterio de conducta específica provee un estándar

absoluto con respecto del cual se compara el logro de aprendizaje de un individuo. Mencionando también a Robert Glaser como fuente de esta propuesta, Norman Gronlund indica que en virtud de que es este estándar de criterio más que la posición relativa en un grupo de norma el que es empleado para describir el rendimiento en una prueba, estas interpretaciones son denominadas referidas a criterios (pág. 2). La precisión requerida para la definición de criterios en términos de conductas específicas condujo a postular que los objetivos de aprendizaje debían ser formulados de manera operacional, postura que tuvo cierta dominancia durante algún tiempo en la práctica de la programación didáctica. Sin embargo, y a pesar de esta clara diferenciación que se establece entre los dos tipos de evaluación, Perrenoud (1990) los asocia, entendiendo que ambos están presentes en la evaluación formal:

> Las prácticas observables participan de una combinación intuitiva de la evaluación de referencia normativa y de la evaluación de referencia criterial, ambas muy artesanales, cada una de las cuales neutraliza en parte los posibles excesos de la otra (pág. 132).

En el enfoque de la evaluación referida a normas en la que se emplean los puntajes derivados, los tipos de normas de uso más frecuente son los rangos percentiles, las normas de edad y de grado y las normas de puntaje estándar.

1. *Rangos percentiles*: el percentil representa el porcentaje de todos los puntajes brutos que se hallan por debajo de un puntaje bruto dado. La escala de rangos percentiles se confecciona ordenando en forma decreciente los puntajes brutos obtenidos por los alumnos al ser administrada la prueba de evaluación. Se construye una tabla de distribución de frecuencias, anotando junto a cada intervalo de puntaje el número de alumnos que ha obtenido dicho puntaje. Este número se denomina frecuencia (f). Junto a la columna f se ubica la columna de las frecuencias acumuladas que se determina sumando de abajo hacia arriba las frecuencias de los intervalos de puntaje.

El cálculo de los percentiles requiere la aplicación de una fórmula que permite tener en cuenta cuál es el puesto de cada alumno en el conjunto total de la clase. Si se encuentra entre los percentiles 90 a 99,

un estudiante demuestra tener un puntaje superior al del 90 a 99 % de los compañeros que se han tomado como referencia para la construcción de la escala de percentiles. Si se encuentra en el percentil 50, esto significa que la mitad de sus compañeros tienen un puntaje inferior al de ese alumno, lo cual equivale a la mediana. Si se ubica en el percentil 20, su puntaje supera sólo al del 20 % de sus compañeros. La escala de percentiles está construida, pues, sobre la base de la comparación entre los aprendizajes efectivamente logrados por los estudiantes, y se emplea válidamente para establecer comparaciones, pero no permite medir aprendizajes en términos de niveles deseables de logro.

2. *Normas de edad*: para determinar estas normas se administran pruebas a grupos de alumnos, representativos de varios niveles sucesivos de edad. Se computan luego las medidas de tendencia central (media aritmética, mediana y modo) correspondientes a cada uno de los niveles de edad. De esta manera se establece cuál es el puntaje bruto típico para cada grupo de edad en las pruebas que se han utilizado. Se debe tener presente, precisamente, que esos puntajes sólo se refieren a las pruebas utilizadas y que la extrapolación de los resultados a otras pruebas debe efectuarse con mucha precaución para garantizar que se trata de pruebas razonablemente equivalentes entre sí.

3. *Normas de grado*: tienen características similares en cuanto a su proceso de construcción a las que se mencionaron en el apartado anterior. La diferencia estriba en que aquí se adopta como referencia para la comparación el grupo de grado, año o nivel de escolaridad. Las normas de grado están basadas sobre los puntajes brutos obtenidos al administrar una prueba a grupos representativos de alumnos de grados sucesivos. Se calcula luego la mediana (percentil 50) de los puntajes brutos obtenidos por los alumnos de cada grado o nivel de escolarización. Se construyen así tablas de puntajes que permiten comparar el puntaje de cualquier alumno de ese grado o nivel con el de los alumnos considerados típicos de éste.

4. *Normas de puntaje estándar*: en sus varias expresiones, estas normas se construyen sobre la base de la concepción de que el

rendimiento de los alumnos de un curso o cohorte se distribuye en la forma de una curva normal (curva de Gauss). De esta manera se expresa la idea de que en un grupo hay siempre un pequeño número de alumnos de muy buen rendimiento, un número igualmente pequeño de alumnos de mal rendimiento y un número importante de alumnos que tienen un rendimiento mediano (algo superior, igual o algo inferior a la media del rendimiento de ese grupo). Para determinar cuáles han de ser las normas del puntaje estándar es necesario trabajar a partir de los puntajes brutos calculando la media de los puntajes y el desvío estándar.

Así, es posible conocer el puntaje en torno del cual se concentran la mayoría de los puntajes brutos obtenidos por los alumnos en la prueba y, merced al conocimiento del desvío estándar, cuál es la amplitud de la curva.

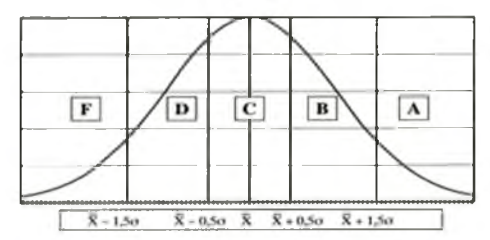
La combinación de estas dos medidas estadísticas permite construir los estándares que se emplearán para obtener los puntajes derivados. Hay tres modalidades que, con escalas diferentes, se emplean con este fin: la "graduación sobre la curva", los "puntajes estándar z" y los "puntajes T".

a) La graduación sobre la curva trabaja habitualmente con una escala de 5 grados en orden de calidad decreciente de los aprendizajes:

A ; B ; C ; D ; F

En esta escala, los primeros cuatro grados (A, B, C y D) son calificaciones de aprobación, en tanto el quinto corresponde a desaprobación o fracaso (F).

Graduación sobre la curva



Como se observa en el gráfico, la banda que corresponde a los puntajes que habrán de recibir la calificación C se encuentra entre +1/2 desvío estándar y -1/2 desvío estándar. La calificación B se asigna a los que se encuentran entre +1/2 y +1 1/2 desvío estándar. Los puntajes que están por encima de +1 1/2 desvío estándar reciben la calificación A. Del mismo modo se procede, pero inversamente, para otorgar los puntajes inferiores a C: a los que se encuentran entre -1/2 y -1 1/2 desvío estándar se les otorga la calificación D y los que están por debajo de -1 1/2 desvío estándar reciben la calificación F de desaprobación.

En esta escala se emplea la medida estadística Desviación Estándar que es la medida de variabilidad más utilizada. La escala combina el empleo de la media como medida de tendencia central con el supuesto de que los datos correspondientes se aproximan a una curva normal. Por las características de esta curva, si los puntajes se distribuyeran de acuerdo con ella, el 34 % de los casos se encontrarían entre la media y un desvío estándar por encima de la media y el 34 % entre la media y un desvío estándar por debajo de ella. En una curva con estas características, dentro de la zona delimitada por tres desviaciones estándar por encima de la media y otras tres por debajo de ella, estarían incluidos todos los casos.

Los "puntajes z" se construyen sobre la base de las mismas medidas estadísticas que la "graduación sobre la curva". Estos

puntajes permiten obtener puntajes combinados provenientes de varias pruebas, homogeneizando el valor de los puntajes aunque cada una de las pruebas emplee una escala de diferente número de grados. Los puntajes z se emplean, en consecuencia, para hacer más sencilla la manipulación de las calificaciones.

La fórmula para obtener puntajes z es la siguiente:

$$\text{Puntaje z} = \frac{\text{Puntaje bruto - Media de los puntajes brutos}}{\text{Desvío Estándar}}$$

Si un alumno, por ejemplo ha obtenido 25 puntos de un total de 40 puntos y la media de los puntajes brutos es de 30 puntos y el desvío estándar es 5:

$$\text{Puntaje z} = \frac{25 - 30}{5}$$

$$\text{Puntaje z} = \frac{-5}{5}$$

$$\text{Puntaje z} = -1$$

Como se ve, si el puntaje del alumno es inferior a la media, su puntaje derivado z será negativo. Si, en cambio, es superior a la media, z será positivo. Cada puntaje z se podrá calcular sobre la base de pruebas cuyo puntaje total es distinto. Si en el caso anterior, el máximo puntaje era 40, en otra prueba podría ser de 25 o de 98. El puntaje z que se obtenga será siempre en unidades positivas o negativas, en enteros o con fracción decimal. La combinación de los puntajes z para definir, por ejemplo, la promoción de un alumno, resultará de la suma algebraica de los puntajes z.

Alumno XC: Puntaje z combinado = +1 +2 + (-3) + 2 + (-1)
= +1

Se puede determinar que la promoción se obtendrá al conseguir

0 punto z o más. La calificación 0 punto z representa que en el conjunto de las instancias de evaluación el alumno se ha colocado en el nivel del promedio de su clase o por encima de éste. La calificación que se establezca para la promoción como puntaje z mínimo deberá resultar de un cuidadoso análisis de los propósitos que se persiguen en el curso o en la escuela y de la importancia que se asigna a la comparación entre los aprendizajes de los alumnos.

Los puntajes T son similares a los anteriores puntajes z, pero se emplean con la intención de no utilizar puntajes positivos y negativos. Se transforma la escala en 0 - 100 puntos, siempre positivos. La fórmula es la siguiente, calculada a partir de los puntajes z de cada alumno:

Puntaje T = (Puntaje z . 10) + 50

Por ejemplo: Alumno FK = (+2 . 10) + 50
= +20 + 50
= 70 puntos T

Alumno JM = (-3 . 10) + 50
= -30 + 50
= 20 puntos T

La escala de puntajes asignados por "graduación sobre la curva" ha sido muy utilizada. También, pero con menor frecuencia, se utilizan las escalas de puntajes z y T.

Como hemos visto, el uso de pruebas estandarizadas se apoya, fundamentalmente, en la necesidad de permitir la comparabilidad de las calificaciones y garantizar su confiabilidad. Pero, como indica Sachs Adams (1970), "no se puede aplicar mecánicamente este procedimiento de calificación en todos los cursos, debido a las diferencias existentes entre los resultados de los diversos cursos, la labor desarrollada en cada uno de ellos, etcétera." Y agrega: "No se pueden evitar preguntas como la siguiente: un curso concreto ¿puede considerarse representativo de la población escolar general?" (pág. 676). Después de analizar diferentes efectos de la adopción de grupos más o menos numerosos o amplios, por un lado, y de grupos más específicos, por el otro, concluye que "1) no es lógico calificar a cada curso según una política de calificaciones uniforme para toda

la escuela; 2) habrá que modificar la política de calificación para cursos determinados sobre la base de las diferencias existentes entre los diferentes cursos con respecto a uno o varios tests correlacionados a las diferencias de competencia individual con respecto a la materia en cuestión" (pág. 680), ya que esas diferencias son aleatorias y si se ignoraran podrían influir de modo decisivo e injusto en las calificaciones de los alumnos menos dotados por comparación con su grupo de compañeros de aula o del resto de la escuela.

Aun cuando se han reconocido ventajas en el uso de pruebas estandarizadas para el cumplimiento de alguna de las funciones de la evaluación, han recibido fuertes críticas, además de la mencionada, en razón de que es innegable que contribuyen a acrecentar la competencia y el conflicto entre los estudiantes porque las altas calificaciones de los alumnos de mejor puntaje bruto incrementan el puntaje de la media y, por tanto, aumentan las posibilidades de desaprobación de un mayor número de estudiantes. Otra crítica de que han sido objeto proviene de la comprobación de que, en algunos casos, por acuerdo entre los alumnos se ha disminuido el puntaje bruto de la prueba intencionalmente para que, al ser inferior la media, los puntajes requeridos para la aprobación resultaran igualmente inferiores. Pero, en última instancia, es el supuesto básico sobre el que se construye esta escala el que merece las críticas más negativas porque es incorrecta la concepción de que los puntajes, y las personas según la calidad de sus aprendizajes o en relación con sus talentos, se distribuyen en la forma de una curva normal. En primer lugar, porque ésta no es una ley de la naturaleza. Los distintos aspectos que se evalúan pueden distribuirse y, por tanto, representarse con diversos tipos de curvas. Si un instrumento de evaluación se utiliza para evaluar los conocimientos básicos que todos los alumnos deben haber aprendido, la curva debería tener una amplitud cero. Y respecto de los talentos, en realidad se puede encontrar mucha gente altamente dotada y, de hecho, esto dependerá también del umbral que se considere punto de partida en cada caso. Las diferencias, de ese modo, pueden aparentar ser pequeñas o grandes, lo cual sería producto de una combinación del efecto umbral y de la escala que en cada caso se utilizara.

Evaluación de portafolio

Se trata de una evaluación que se realiza sobre las producciones de los alumnos. Al modo en que un profesional presenta su trabajo y se presenta a sí mismo en una situación de búsqueda de empleo o similar, el alumno selecciona sus mejores trabajos y los pone a disposición del análisis del docente. Para el docente y para el alumno, esta carpeta o portafolio de trabajos tiene la función de representar los aprendizajes que el alumno ha realizado a lo largo de su curso. Esta estrategia de evaluación se apoya sobre varias ideas: la evaluación de certificación o promoción no se efectúa sobre la base del desempeño del alumno en una instancia desarrollada en un lapso breve. Los trabajos son los que se han ido produciendo durante el transcurso de todo el curso. El alumno hace una evaluación de sus propios trabajos, debiendo analizarlos en sus logros, sus errores y sus dificultades. La forma de presentación de los trabajos puede ser variada, incluyéndose producciones escritas, gráficas, grabadas, videos, objetos, etcétera. El alumno puede mostrar la evolución de su desempeño, su mejoramiento progresivo, discutir los problemas y aspectos que no ha podido resolver todavía. Quienes utilizan esta modalidad de evaluación señalan la importancia pedagógica de que la evaluación en una situación no genera los niveles de ansiedad típicos de las situaciones de examen, el valor que se da a la responsabilidad del alumno, su papel en la puesta en juego de los mecanismos de autoevaluación que acompañan a todo el proceso de producción y la autoestima que contribuye a desarrollar en el alumno al reconocerlo como participante necesario y no como objeto del proceso de evaluación. El logro de estas características positivas depende de la actitud con la que el docente asuma su función evaluadora y de la relación que, en consecuencia, entable con sus alumnos. Como forma única de decidir la certificación de los aprendizajes o como complementaria de otras modalidades, la evaluación de portafolio es insustituible a la hora de evaluar la producción de los alumnos. A las otras ventajas que ya mencionamos debemos agregar que es un factor determinante en la programación de actividades de aprendizaje, ya que requiere que los alumnos cuenten con trabajos realizados por ellos, esto es, con una carpeta o portafolio de trabajos, lo cual estimula la adopción de un

modo de enseñar y de aprender centrados en tareas productivas, no rutinarias. Esto por sí mismo es bueno desde la perspectiva de la formación que la escuela debe brindar.

La promoción y los juicios de valor

Hemos desarrollado los principios básicos de seis modelos de régimen de promoción. Son los que presentan mayor frecuencia de administración en las instituciones escolares.

Un problema que ha dado lugar a controversias es el de la importancia otorgada a la cuestión de la comparabilidad. Diversos autores adoptan una postura crítica ante el establecimiento de sistemas de calificación y regímenes de promoción en los que la comparación entre alumnos juega un relevante papel. Entre ellos, Cardinet (1989) señala las diferencias esenciales que surgen de la contrastación de los tests psicométricos que sirvieron en su momento de modelo a las pruebas formales de evaluación, y estas últimas cuando se administran con sentido pedagógico. En los tests psicométricos se procura elaborar un diagnóstico exhaustivo pero con el mínimo número de datos necesarios (por aplicación del principio de parsimonia). Se buscan, también, características estables, estandarizadas, y las escalas cuantitativas que se emplean son relativas a los estándares "normales" para la población estudiada. En pedagogía, en cambio –sostiene Cardinet–, la búsqueda de la parsimonia es totalmente inexistente. Todos los contenidos son esenciales. Una habilidad ausente no puede compensarse con la presencia de otra habilidad. Todas las combinaciones posibles de contenidos, de procedimientos cognitivos y de situaciones de observación, constituyen objetivos independientes e irreductibles. "El deseo de obtener características estables –dice Cardinet– está muy ausente en la escuela" (pág. 45 y sig.). Las escalas cuantitativas son, además, más bien absolutas que relativas. El papel de la situación es, finalmente, de enorme importancia para que el evaluador pueda comprender el significado de la conducta del alumno. Acentúa Cardinet, por otra parte, la necesidad de incluir en el análisis la relación del evaluador y el evaluado como un aspecto determinante de esa situación. Ante las características que son propias del acto de

evaluar, según este autor, la evaluación pedagógica debe evitar la emisión de juicios que serían incompatibles con la ayuda que el docente debe dar a su alumno. Éste sí necesita información sobre su desempeño pero, es la conclusión de Cardinet, sin juicios de valor.

Queremos resaltar, sin embargo, que el juicio de valor que Cardinet rechaza es, básicamente, el que resulta de la comparación entre diferentes alumnos. Queremos pensar, en cambio, en una evaluación que resguarda todas las consideraciones que hemos ido mencionando hasta aquí, que conserva la conciencia de la precariedad de los juicios evaluativos, en la que el evaluador asume el modesto papel que está implicado en el acto de medir aspectos de la conducta humana, en donde una persona evalúa a otra y hay evaluaciones mutuas en el curso de la interacción. Esta evaluación lejos de perder su papel pedagógico lo asume plenamente y se estructura como componente esencial de los actos de enseñanza y de aprendizaje. Para ello es menester que el docente integre a su labor de enseñanza la tarea de la evaluación: "Un sistema de registro acumulativo racional es indispensable –afirma Sachs Adams (1970)– para permitir que la medición y la evaluación alcancen una eficacia máxima para facilitar la comprensión de los alumnos y de sus problemas" (pág. 682). El registro no es suficiente, sin embargo. Hay que analizar e interpretar las informaciones.

La promoción se resuelve siempre a partir de una decisión y ella lleva implícito un juicio de valor. El régimen de promoción que se establezca debe dar a este acto la mayor explicitación posible, de modo de garantizar la transparencia y la justicia de la decisión.

Bibliografía

Cardinet, Jean: "Evaluer sans juger", *Révue Française de Pédagogie*, nº 88, julio-agosto-setiembre de 1989.

De Ketele, Jean-Marie: *Observar para educar*, Madrid, Aprendizaje-Visor, 1984.

—: "L'évaluation conjuguée en paradigmes", *Révue Française de Pédagogie*, nº 103, abril-mayo-junio de 1993.

Gronlund, Norman E.: *Preparing Criterion-Referenced Tests for Classroom Instruction*, The Macmillan Co., Collier-Macmillan Publishers, 1973.

Hotyat, E.: *Los exámenes*, Buenos Aires, Kapelusz, 1965.

House, E. R.: *Evaluación, ética y poder*, Madrid, Morata, 1994.

Lavin, David E.: *The Prediction of Academic Performance*, Russell Sage Foundation, 1965.

Perrenoud, P.: *La construcción del éxito y el fracaso escolar*, Madrid, Paideia-Morata, 1990.

Piobetta, J. B.: *Exámenes y concursos*, Buenos Aires, Kapelusz, 1952.

Popham, W. James: *Evaluación basada en criterios*, Madrid, Magisterio Español, 1983.

Purkey, William W.: *Self Concept and School Achievement*, Prentice-Hall Inc., 1970.

Sachs Adams, Georgia: *Medición y evaluación en educación, psicología y guidance*, Barcelona, Herder, 1970.

Thorndike, Robert E. y Hagen, Elizabeth P.: *Medición y evaluación en psicología y en educación*, México, Trillas, 1991.

Wiliam, Dylan: "Some technical issues in assessment: a user's guide", *British Journal for Curriculum and Assessment*, 2 (3), 1992.